

NUMBER OF SEQ ID NOS: 10535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9696
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
US-09-649-163-9696

Query Match 98.5% Score 1548.8 DB 25 Length 2494
Best Local Similarity 99.5% Pred. No. 2.9e-298
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
QY 1 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACAGCAGCATAGAGACTGAAGCAG 60
DB 882 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACAGCAGCATAGAGACTGAAGCAG 941
QY 61 ACCCCGCGCCAGCGGAGCG -CCCCACAGCCAGGAGACCCCTGGCAGCGGAGCGCG 119
DB 942 ACCCCGCGCCAGCGGAGCGGCCCCACAGCGGAGACCCCTGGCAGCGGAGCGCG 1001
QY 120 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 179
DB 1002 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 1061
QY 180 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 239
DB 1062 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 1121
QY 240 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 299
DB 1122 GGTGAGGTTATGATCCAGCGGGGGGCCCGGGGGGCTGCTCCCGGGCCCTGGCGCGT 1181
QY 300 GAACCAAGCGGGGAGCTGGTGGGCTGGAGGAGCTGGGCGCTGGAGGCTGGTGGT 359
DB 1182 GAACCAAGCGGGGAGCTGGTGGGCTGGAGGAGCTGGGCGCTGGAGGCTGGTGGT 1241
QY 360 CATGCTGAGAGAGGAGCTGATGACAGAGGTGAGAGGAGCTGATGAGCGGCTGACTG 419
DB 1242 CATGCTGAGAGAGGAGCTGATGACAGAGGTGAGAGGAGCTGATGAGCGGCTGACTG 1301
QY 420 GAGAGCGGCTGATGACAGAGGCTGATGAGAGGCTGATGAGAGGCTGATGAGAGG 479
DB 1302 GAGAGCGGCTGATGACAGAGGCTGATGAGAGGCTGATGAGAGGCTGATGAGAGG 1361
QY 480 AGCTGCTGAACCATTAATGCTGATGAGAGGCTGATGAGAGGCTGATGAGAGGCT 539
DB 1362 AGCTGCTGAACCATTAATGCTGATGAGAGGCTGATGAGAGGCTGATGAGAGGCT 1421
QY 540 CTGACAGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
DB 1422 CTGACAGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
QY 600 GGTGCTGGGGCTGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
DB 1482 GGTGCTGGGGCTGCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541
QY 660 GAGCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
DB 1542 GAGCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1601
QY 720 CTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
DB 1602 CTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1661
QY 780 AGTGGGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
DB 1662 AGTGGGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1721
QY 840 CTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
DB 1722 CTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1781
QY 900 TGTGCTAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
```

```
DB 1782 TGTGCTAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1841
QY 960 GGGCCGCTGTCAGAGCTGGGCTGATGATGTTTACGTGCGGCGGAGAGTCTGTCG 1019
DB 1842 GGGCCGCTGTCAGAGCTGGGCTGATGATGTTTACGTGCGGCGGAGAGTCTGTCG 1901
QY 1020 CATGCTGCTGCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
DB 1902 CATGCTGCTGCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1961
QY 1080 CTACTTGTATATGTCGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
DB 1962 CTACTTGTATATGTCGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2021
QY 1140 GTTTCAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
DB 2022 GTTTCAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2081
QY 1200 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 2082 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2141
QY 1260 GCCACCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
DB 2142 GCCACCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2201
QY 1320 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1379
DB 2202 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2261
QY 1380 GAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1439
DB 2262 GAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2321
QY 1440 TGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
DB 2322 TGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2481
QY 1500 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559
DB 2382 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441
QY 1560 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1571
DB 2442 CACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2453
```

RESULT 9
US-09-652-126-10026
Sequence 10026, Application US/09652126
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 1600.1185-001
CURRENT APPLICATION NUMBER: US/09/652,126
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/151,132
NUMBER OF SEQ ID NOS: 10051
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10026
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
US-09-652-126-10026

Query Match 98.5% Score 1548.8 DB 25 Length 2494
Best Local Similarity 99.5% Pred. No. 2.9e-298
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 GCCCCACAGCCGCGCCCTGGCAGCCCGCTGGGACAGCAGCATAGAGACTGAAGCAG 60

```
Db 882 GGGCCACACACCGGCTCTGACAGCCGCTGGGACACACCGATTAGAGAGCTGAAAGGAGG 941
Qy 61 AGCCGCCCCCAGGGGACGG -CCCCACAGCGGCCAGGGAGCCCTCTGGCAGCGGAGCCGCG 119
Db 942 AGCGGCGCCAGCGGACGCCGCCACAGGCGCCAGGACCCCTTGGCAGCGGAGCGCGG 1001
Qy 120 GGTGAGGTTATGATGATCCAGCGGCGGCGCCCGCGGCGGTGCTCCCGGCGCTGCGCGT 179
Db 1002 GGTGAGGTTATGATGATCCAGCGGCGGCGCCCGCGGCGGTGCTCCCGGCGCTGCGCGT 1061
Qy 180 GGTGAGGTTATGATGATCCAGCGGCGGCGCCCGCGGCGGTGCTCCCGGAGTCA 239
Db 1062 GGTGAGGTTATGATGATCCAGCGGCGGCGCCAGGAGGCGCTTGGAGCTCTTCCGAGTCA 1121
Qy 240 GGTGAGGTTATGATGATCCAGCGGCGGCGCTTGGAGCTCTTGGAGCTCTTGGAGCTCA 299
Db 1122 GGTGAGGTTATGATGATCCAGCGGCGGCGCTTGGAGCTCTTGGAGCTCTTGGAGCTCA 1181
Qy 300 GAACACAGCGCGGAGCTGTGCTGCGAGAGCTGTGCGCGCTGTGCGAGCTGTGCTGTG 359
Db 1182 GAACACAGCGCGGAGCTGTGCTGCGAGAGCTGTGCGCGCTGTGCGAGCTGTGCTGTG 1241
Qy 360 CATGTGTGAGAGCGGCGTGTGATGACAGAGTGTGTAACGGGCTCATGAGGCGGCTGACTG 419
Db 1242 CATGTGTGAGAGCGGCGTGTGATGACAGAGTGTGTAACGGGCTCATGAGGCGGCTGACTG 1301
Qy 420 GAGAGCGGCGCATCCAGAGCGGCGCTGTGAGCTGTGAGAGCGGCTGTGAGAGCGGCTG 479
Db 1302 GAGAGCGGCGCATCCAGAGCGGCGCTGTGAGCTGTGAGAGCGGCTGTGAGAGCGGCTG 1361
Qy 480 AGCTTCTTGAACCATTTATGCTGTGATGAGAGCTGTGACCAATGAAGACCTGTGACCA 539
Db 1362 AGCTTCTTGAACCATTTATGCTGTGATGAGAGCTGTGACCAATGAAGACCTGTGACCA 1421
Qy 540 CTGACAGCTATTTGCTGTGCTGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 599
Db 1422 CTGACAGCTATTTGCTGTGCTGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
Qy 600 GGGCTTGGGGGCTGCGGCTCTTCTGTGTGCTGACGCTGCGGCTGCGGCTTCAATGCTGAT 659
Db 1482 GGGCTTGGGGGCTGCGGCTCTTCTGTGTGCTGACGCTGCGGCTGCGGCTTCAATGCTGAT 1541
Qy 660 GGAAGCTAGAGATGAGAGATGTCGCGCTGTGCGGAGATGAGCTTCACTGTGCGACCTT 719
Db 1542 GGAAGCTAGAGATGAGAGATGTCGCGCTGTGCGGAGATGAGCTTCACTGTGCGACCTT 1601
Qy 720 CCGAGCTGTGAGAGCTGTGCGGCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCTG 779
Db 1602 CCGAGCTGTGAGAGCTGTGCGGCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCTG 1661
Qy 780 AGTGGGTTCCAGAGACCTGCTGCGGCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCT 839
Db 1662 AGTGGGTTCCAGAGACCTGCTGCGGCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCT 1721
Qy 840 CTTTGTGCTGCTGAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
Db 1722 CTTTGTGCTGCTGAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1781
Qy 900 TGTGCTAGTCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Db 1782 TGTGCTAGTCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1841
Qy 960 GGGGCTGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
Db 1842 GGGGCTGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1901
Qy 1020 CATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
Db 1902 CATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1961
Qy 1080 CTACTTGTGATATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
```

```
Db 1962 CTACTTGTGATATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2021
Qy 1140 GTTTCAGTGTGATGCGGCAATTTGATGTTAGCAGCGCTGTGAGCGGCGGCTGACCCAAA 1199
Db 2022 GTTTCAGTGTGATGCGGCAATTTGATGTTAGCAGCGCGCTGTGAGCGGCGGCTGACCCAAA 2081
Qy 1200 CTACTTGTGATGCTGACGCGGTTGCTGTGAGCGCGCGGCGGCTGAGAGCGGCAAGAT 1259
Db 2082 CTACTTGTGATGCTGACGCGGTTGCTGTGAGCGCGCGGCGGCTGAGAGCGGCAAGAT 2141
Qy 1260 GCGACCGCGAGAGAGCGGCTTATGAGCGGCTGCGCGGCTGCGGCTGCTGCTGCTGCTG 1319
Db 2142 GCGACCGCGAGAGAGCGGCTTATGAGCGGCTGCGCGGCTGCGGCTGCTGCTGCTGCTG 2201
Qy 1320 AGGACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
Db 2202 AGGACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2261
Qy 1380 GAGACTCTGCTGAGAGAGGCTGAGAGGCTGATGCTTTGGGGGAGAGGCGAGAA 1439
Db 2262 GAGACTCTGCTGAGAGAGGCTGAGAGGCTGATGCTTTGGGGGAGAGGCGAGAA 2321
Qy 1440 TGAAGTCTGCTGAGAGGCGGAGGCGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTG 1499
Db 2322 TGAAGTCTGCTGAGAGGCGGAGGCGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTG 2381
Qy 1500 TTTTGTTCGACACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1559
Db 2382 TTTTGTTCGACACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2441
Qy 1560 AAAAAAAAAA 1571
Db 2442 CACAACATGAA 2453

RESULT 10
US-09-937-10528
; Sequence 10528, Application US/09652913
; GENERAL INFORMATION:
; APPLICANT: Faib, Dean R.
; TITLE OF INVENTION: NOVEL, NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1174-001
; CURRENT APPLICATION NUMBER: US/09/652,913
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ. ID NOS.: 10833
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ. ID NO. 10528
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-913-10528

Query Match 98.5%; Score 1548.8; DB 25; Length 2494;
Best Local Similarity 99.5%; Pred. No. 2,96-298;
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
```

Db	1062	GCAGTGTGTGTAAACCCCGCGGGGCAAAAGGCAAGGCTTGCAGCTCTTCGGAGTCA	1122
Oy	240	CGTGCAGCCCTTTTGGCTGAGGCTGTAAATCTTCCTTCAACGCTGATGCTCACTGAGCGGG	299
Db	1122	CGTGCAGCCCTTTTGGCTGAGGCTGTAAATCTTCCTTCAACGCTGATGCTCACTGAGCGGG	1181
Oy	300	GAACCAAGCGCGGGAGCTGTGGGGTCCGAGAGAGCTGGGGCGGCTGGAGAGCTGTGGTGGT	359
Db	1182	GAACCAAGCGCGGGAGCTGTGGGGTCCGAGAGAGCTGGGGCGGCTGGAGAGCTGTGGTGGT	1241
Oy	360	CATGCTGTGAGACGGGCTGTATGTCACAGAGGTGGGAACGGGCTCATGTAGAGCGGCTGACGTG	419
Db	1242	CATGCTGTGAGAGCGGCTGTATGTCACAGAGGTGGGAACGGGCTCATGTAGAGCGGCTGACGTG	1301
Oy	420	GGAGACCGCCATCCAGAAAGCCCTGTGTAGCTCCAGAGGAGCTCTGGCAACGACCGTGGC	479
Db	1302	GGAGACCGCCATCCAGAAAGCCCTGTGTAGCTCCAGAGGAGCTCTGGCAACGACCGTGGC	1361
Oy	480	AGCTTCTTGAACATTATATGCTGGCTATATAGCAAGGTCAACCAATGAAGACCTCTGACCA	539
Db	1362	AGCTTCTTGAACATTATATGCTGGCTATATAGCAAGGTCAACCAATGAAGACCTCTGACCA	1421
Oy	540	CTGCACGCTATTTGTGTGGCGCGGGGCTGTACCCATGAACCTGTGTCTGTCAACAC	599
Db	1422	CTGCACGCTATTTGTGTGGCGCGGGGCTGTACCCATGAACCTGTGTCTGTCAACAC	1481
Oy	600	GGCTTTCGGGGCTGGCCCTCTTCTGTGTGTCAACCTGGCGCTGGGCTTCATTATGCTGATGT	659
Db	1482	GGCTTTCGGGGCTGGCCCTCTTCTGTGTGTCAACCTGGCGCTGGGCTTCATTATGCTGATGT	1541
Oy	660	GGACCTAGAGAGTGAAGATATCGCGCTGTGGGGAGATGCGCTTCACTCTGGGACCTT	719
Db	1542	GGACCTAGAGAGTGAAGATATCGCGCTGTGGGGAGATGCGCTTCACTCTGGGACCTT	1601
Oy	720	CCTGGTGTGGACACCTCGGGGCACTAACGGGGCGAGTGGGCTCACTCCCTGTAGAGAG	779
Db	1602	CCTGGTGTGGACACCTCGGGGCACTAACGGGGCGAGTGGGCTCACTCCCTGTAGAGAG	1661
Oy	780	AGTGGGTCCAAAGACACCTCCCTCCCGCTTGTGTGTGTGCACAGGGGCGCGGTAGATGCACA	839
Db	1662	AGTGGGTCCAAAGACACCTCCCTCCCGCTTGTGTGTGTGCACAGGGGCGCGGTAGATGCACA	1721
Oy	840	CCTTGTGCACTGGAGAGCCAGTGGCCCTCTCACTGGACAGTGGTCCCGACGAGGACTT	899
Db	1722	CCTTGTGCACTGGAGAGCCAGTGGCCCTCTCACTGGACAGTGGTCCCGACGAGGACTT	1781
Oy	900	TGTGTATCTCTGGCACTGCGTGCACATCCGCACTGGGCACTGAGATGTTTGTGTGCACCAT	959
Db	1782	TGTGTATCTCTGGCACTGCGTGCACATCCGCACTGGGCACTGAGATGTTTGTGTGCACCAT	1841
Oy	960	GGGGCGCTGTCACTGCGCTCATATGCATCTGTTCTTACGTGGCGGCGGAGTGTCTGTGTC	1019
Db	1842	GGGGCGCTGTCACTGCGCTCATATGCATCTGTTCTTACGTGGCGGCGGAGTGTCTGTGTC	1901
Oy	1020	CATGCTGTGCGCTCTTCTTGCCCATGAGAAAGGCGAGGCAATATGAGATATGAATGCCC	1079
Db	1902	CATGCTGTGCGCTCTTCTTGCCCATGAGAAAGGCGAGGCAATATGAGATATGAATGCCC	1961
Oy	1080	CTACTGTGTATATGTGCCCGCTGTGTCCGCTTCCGCTTGGAGACCCCAAGATGGGAAGGTGT	1139
Db	1962	CTACTGTGTATATGTGCCCGCTGTGTCCGCTTCCGCTTGGAGACCCCAAGATGGGAAGGTGT	2021
Oy	1140	GTTTGCAGTGTATGGGAATTTGATGTGTAGCCAGGCGCGTGCAGGGGCAAGTGCACCCCAA	1199
Db	2022	GTTTGCAGTGTATGGGAATTTGATGTGTGTAGCCAGGCGCGTGCAGGGGCAAGTGCACCCCAA	2081
Oy	1200	CTACTTCTGATGTCAAGCGGTTTCCGTGTGAAGCCCGCCGACCTGTGAAGCCCCACAGAT	1259
Db	2082	CTACTTCTGATGTGTCAAGCGGTTTCCGTGTGAAGCCCGCCGACCTGTGAAGCCCCACAGAT	2141
Oy	1260	GGCAACGGCAAAAGAGCCCTTATAGACCCCTGGGCGCGCTGTGTCTTAACTGTCTACTTGC	1319
Db	2142	GGCAACGGCAAAAGAGCCCTTATAGACCCCTGGGCGCGCGCTGTGTCTTAACTGTCTACTTGC	2201

```

QY 1320 AGACCCCTTCCTCCCTTCCTTAGGGCTGCAGGGGCTCTCCACAGCTCCTTCGGGGGTGGAG 1379
Db 2202 AGAGCCCTTCCTCCCTTCCTTAGGGCTGCAGGGGCTCTCCACAGCTCCTTCGGGGGTGGAG 2261
QY 1380 GAGACTCCTCTGCAGAAAGGTGAGAAAGTGGAGGGCTATCTTTGGGGGAGACAGCCAGAA 1439
Db 2262 GAGACTCCTCTGCAGAAAGGTGAGAAAGTGGAGGGCTATCTTTGGGGGAGACAGCCAGAA 2321
QY 1440 TGAAGTCTGAGACCCGCCACCCGACGACCAACCAATTCAGAAATGAAGTGACATTCACAGCTGA 1499
Db 2322 TGAAGTCTGAGACCCGCCACCCGACGCTGGCTGGGGCCAGCTGCTATGTAAAGCCTTCAG 2381
QY 1500 TTTGTTCTGAGACCCGCCACCCGACGACCAACCAATTCAGAAATGAAGTGACATTCACAGCTGA 1559
Db 2382 TTTGTTCTGAGACCCGCCACCCGACGACCAACCAATTCAGAAATGAAGTGACATTCACAGCTGA 2441
QY 1560 AAAAAAAAAAAAAA 1571
Db 2442 CACAAACATGAA 2453

RESULT 11
US-09-652-918-8605
; Sequence 8605, Application US/09652918
; GENERAL INFORMATION:
; APPLICANT: Galvin, Katherine
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1187-001
; CURRENT APPLICATION NUMBER: US/09/652, 918
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,130
; NUMBER OF SEQ ID NOS: 8985
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8605
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-918-8605

Query Match 98.5%; Score 1548.8; DB:25; Length 2494;
Best Local Similarity 99.5%; Pred.No.2.9e-298;
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1

QY 1 GCCCCACACCGCGCCCTGCGAGCGCCGCTGGGCAAGCACCGATTAAGAGCTGAAGGACAG 60
Db 882 GCCCCACACCGCGCCCTGCGAGCGCCGCTGGGCAAGCACCGATTAAGAGCTGAAGGACAG 941
QY 61 AGCCCGCCGACAGGGGACGGC-CGCCACAGCGCCAGGAGCCCTTCGACCGGGAGCCGG 119
Db 942 AGCCCGCCGACAGGGGACGGC-CGCCACAGCGCCAGGAGCCCTTCGACCGGGAGCCGG 1001
QY 120 GGTGCGAGTTATGATCCAGCGGGGCGGCCCGGGGGCGTGCCTCCGGGGCCCGGCCGGT 179
Db 1002 GGTGCGAGTTATGATCCAGCGGGGCGGCCCGGGGGCGTGCCTCCGGGGCCCGGCCGGT 1061
QY 180 GCTGTGCTGTGAACCCGCGGGGCGGCAAGGGCAAGGCTTTCGACAGCTTTCGGAGTCA 239
Db 1062 GCTGTGCTGTGAACCCGCGGGGCGGCAAGGGCAAGGCTTTCGACAGCTTTCGGAGTCA 1121
QY 240 CGTGCAGCCCTTTTGGGTGCTGAGGCTGAAATTCCTTCACGCTGATGCTACATGAGCGGC 299
Db 1122 CGTGCAGCCCTTTTGGGTGCTGAGGCTGAAATTCCTTCACGCTGATGCTACATGAGCGGC 1181
QY 300 GAACACAGCGCGGAGAGCTGCTGGGTGCGAGAGAGCTGGGCGCTGGGAGCGCTGGTGGT 359
Db 1182 GAACACAGCGCGGAGAGCTGCTGGGTGCGAGAGAGCTGGGCGCTGGGAGCGCTGGTGGT 1241
QY 360 CATGTCGAGAGCGGGCTGATGCACAGGCTGTTGAACGGGCTCATGAGACGGGCTGACATG 419

```


QY 600 GCGTTCGGGGGCTGCGCCTCTCTCTGTCGTCAGACCTGGCCCTGGGGCTTCATTTGATGT 659
 DB 1482 GCGTTCGGGGGCTGCGCCTCTCTCTGTCGTCAGACCTGGCCCTGGGGCTTCATTTGATGT 1541
 QY 660 GACCTAGAGAGTGAAGATATCGGGGCTGCGGGGAGATGCGGTTCACCTTGGGACCTT 719
 DB 1542 GACCTAGAGAGTGAAGATATCGGGGCTGCGGGGAGATGCGGTTCACCTTGGGACCTT 1601
 QY 720 CCGGCTTCGGGAGCGCCCTGCGACCTACCGCGCGGACCTGGCCCTACCTCTAGGAAG 779
 DB 1602 CCGGCTTCGGGAGCGCCCTGCGACCTACCGCGCGGACCTGGCCCTACCTCTAGGAAG 1661
 QY 780 AGTGGTTCGAAGACACCTGCGCCTCCCGCTTGTGTCACAGACGCGCGGTAGATGACA 839
 DB 1662 AGTGGTTCGAAGACACCTGCGCCTCCCGCTTGTGTCACAGACGCGCGGTAGATGACA 1721
 QY 840 CCTTGTGCGACCTGAGAGGACGACCTGCGCTTCTACTGACACAGTGGTCCCGACAGACCTT 899
 DB 1722 CCTTGTGCGACCTGAGAGGACGACCTGCGCTTCTACTGACACAGTGGTCCCGACAGACCTT 1781
 QY 900 TGTGCTAGTCCCTGCGACCTGCTGACCTGCGACCTGGGCGAGTGGTCTGACCCAT 959
 DB 1782 TGTGCTAGTCCCTGCGACCTGCTGACCTGCGACCTGGGCGAGTGGTCTGACCCAT 1841
 QY 960 GGGCCGCTGTGACGCTGCGCTCATGATCTGTTCTACGTGCGGGCGGAGTCTCGTGC 1019
 DB 1842 GGGCCGCTGTGACGCTGCGCTCATGATCTGTTCTACGTGCGGGCGGAGTCTCGTGC 1901
 QY 1020 CATGCTGCTGCGCCTCTCTCTGCGCATGAGAAAGGCGACGATATGAGATGATATCCC 1079
 DB 1902 CATGCTGCTGCGCCTCTCTCTGCGCATGAGAAAGGCGACGATATGAGATGATATCCC 1961
 QY 1080 CTACTGTGATATGTCGCGCTGCTGCGCTTCCGCTTGGACGCCCAAGATGGAAAGTGT 1139
 DB 1962 CTACTGTGATATGTCGCGCTGCTGCGCTTCCGCTTGGACGCCCAAGATGGAAAGTGT 2021
 QY 1140 GTTTCGAGTGGATGGGAATTGATGTTTACGAGGCGGTGACGAGCGGCGGACGCCAAA 1199
 DB 2022 GTTTCGAGTGGATGGGAATTGATGTTTACGAGGCGGTGACGAGCGGCGGACGCCAAA 2081
 QY 1200 CTACTGTGATATGTCGCGCTGCTGCGCTTCCGCTTGGACGCCCAAGATGGAAAGTGT 1259
 DB 2082 CTACTGTGATATGTCGCGCTGCTGCGCTTCCGCTTGGACGCCCAAGATGGAAAGTGT 2141
 QY 1260 GCGACGCGCGAGAGAGCGCTTATGACCCCTGGGCGCGCTGCTGCTGCTACTTGC 1319
 DB 2142 GCGACGCGCGAGAGAGCGCTTATGACCCCTGGGCGCGCTGCTGCTGCTACTTGC 2201
 QY 1320 AGGACCCCTTCTCTCTCTAGGCGCTGACAGGCGCTGTCACAGCTCCTGTGGGGGTGAG 1379
 DB 2202 AGGACCCCTTCTCTCTCTAGGCGCTGACAGGCGCTGTCACAGCTCCTGTGGGGGTGAG 2261
 QY 1380 GAGACTCTCTGAGAGAGGCTGAGAGGCTATGCTTTGGGGGACAGGCCGAA 1439
 DB 2262 GAGACTCTCTGAGAGAGGCTGAGAGGCTATGCTTTGGGGGACAGGCCGAA 2321
 QY 1440 TGAAGTCTGAGAGAGCGCTGAGAGGCTGAGAGGCTGAGAGGCTTCTAG 1499
 DB 2322 TGAAGTCTGAGAGAGCGCTGAGAGGCTGAGAGGCTGAGAGGCTTCTAG 2381
 QY 1500 TTTGTCGAGAGCGCGACCCGACGAAACAATCAATTAAGTGAATTCAGAGCTGA 1559
 DB 2382 TTTGTCGAGAGCGCGACCCGACGAAACAATCAATTAAGTGAATTCAGAGCTGA 2411
 QY 1560 AAAAAAAAAA 1571
 DB 2442 CACAAACATGAA 2453

: APPLICANT: Holtzman, Douglas A.
 : APPLICANT: Gearing, David P.
 : TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 : TITLE OF INVENTION: THEREFOR
 : FILE REFERENCE: 1600.2008-001
 : CURRENT APPLICATION NUMBER: US/09/699,998
 : CURRENT FILING DATE: 2000-10-30
 : PRIOR APPLICATION NUMBER: 60/162,362
 : PRIOR FILING DATE: 1999-10-29
 : NUMBER OF SEQ ID NOS: 10905
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 8671
 : LENGTH: 2494
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-699-998-8671
 Query Match 98.5%; Score 1548.8; DB 27; Length 2494;
 Best Local Similarity 99.5%; Pred. No. 2,9e-298;
 Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 QY 1 GCGCCAGAGCGGCGCTGCGAGCGCGCTGGCGACGATTAAGAGTGAAGGAGG 60
 DB 882 GCGCCAGAGCGGCGCTGCGAGCGCGCTGGCGACGATTAAGAGTGAAGGAGG 941
 QY 61 AGCGCGCGCGAGCGAGCG-CCCGACAGCGCGAGGAGCGCGCGAGCGCGCG 119
 DB 942 AGCGCGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1001
 QY 120 GGTGAGAGTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
 DB 1002 GGTGAGAGTATGATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
 QY 180 GGTGAGTCTGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
 DB 1062 GGTGAGTCTGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
 QY 240 CGTGACAGCGCGCTTTGGCTGAGAGTGAATCTCCTTACGCTGATGCTACTGAGCGCG 299
 DB 1122 CGTGACAGCGCGCTTTGGCTGAGAGTGAATCTCCTTACGCTGATGCTACTGAGCGCG 1181
 QY 300 GAACACGCGCGGAGCGTGTGCGGCTGCGAGAGCTGGGCGCTGCTGCTGT 359
 DB 1182 GAACACGCGCGGAGCGTGTGCGGCTGCGAGAGCTGGGCGCTGCTGCTGT 1241
 QY 360 CATGTCGAGAGCGCGCTGATGACGAGAGTGTGAAGCGGCTGTCGAGCGCGCTGACTG 419
 DB 1242 CATGTCGAGAGCGCGCTGATGACGAGAGTGTGAAGCGGCTGTCGAGCGCGCTGACTG 1301
 QY 420 GAGACGCGCATTCAGAGCGCGCTGTAGCTCCAGAGAGCTGTGGCAAGCGCGTGGC 479
 DB 1302 GAGACGCGCATTCAGAGCGCGCTGTAGCTCCAGAGAGCTGTGGCAAGCGCGTGGC 1361
 QY 480 AGCTTCCTTGAACATTATGCTGCTATGAGCAGGTCAACCAATGAAGACCTCTGACCAA 539
 DB 1362 AGCTTCCTTGAACATTATGCTGCTATGAGCAGGTCAACCAATGAAGACCTCTGACCAA 1421
 QY 540 CTGACGCTATGCTGTGCGCGCGCGCGCTGCTGACGAGTCAAGCTGTGCTGCGACAC 599
 DB 1422 CTGACGCTATGCTGTGCGCGCGCGCGCTGCTGACGAGTCAAGCTGTGCTGCGACAC 1481
 QY 600 GCGTTCGGGCGCTGCGCTCTCTGCTGCTGAGCGCTGCGCGCTTCTATGCTGATGT 659
 DB 1482 GCGTTCGGGCGCTGCGCTCTCTGCTGCTGAGCGCTGCGCGCTTCTATGCTGATGT 1541
 QY 660 GACCTAGAGAGTGAAGATATCGGGGCTGCGGGGAGATGCGGTTCACCTTGGGACCTT 719
 DB 1542 GACCTAGAGAGTGAAGATATCGGGGCTGCGGGGAGATGCGGTTCACCTTGGGACCTT 1601
 QY 720 CCTGCGTCTGCGACCGCTGCGCGACCTACCGCGCGGAGATGCGCTTACCTCTGAGAGAG 779
 DB 1602 CCTGCGTCTGCGACCGCTGCGCGACCTACCGCGCGGAGATGCGCTTACCTCTGAGAGAG 1661

Oy 780 AGTGGCTTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 839
Db 1662 AGTGGCTTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1721
Oy 840 CCTGTGCTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 899
Db 1722 CCTGTGCTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1781
Oy 900 TGTGCTAGTCTGAGACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 959
Db 1782 TGTGCTAGTCTGAGACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1841
Oy 960 GGGCGCTGTGAGCTGCGGCTGATCATCTGTTCTACCTCGGCGGAGTGTCTGTGC 1019
Db 1842 GGGCGCTGTGAGCTGCGGCTGATCATCTGTTCTACCTCGGCGGAGTGTCTGTGC 1901
Oy 1020 CATGCTGCTGCGGCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1079
Db 1902 CATGCTGCTGCGGCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1961
Oy 1080 CTACTTGTATGTGCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1139
Db 1962 CTACTTGTATGTGCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 2021
Oy 1140 GTTGTGCTGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1199
Db 2022 GTTGTGCTGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2081
Oy 1200 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
Db 2082 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2141
Oy 1260 GCCACCGCCGAGAGACCCCTTATGACCCCTGCGGCGGCTGTGCTTACTTACTTGC 1319
Db 2142 GCCACCGCCGAGAGACCCCTTATGACCCCTGCGGCGGCTGTGCTTACTTACTTGC 2201
Oy 1320 AGGACCTTCT 1379
Db 2202 AGGACCTTCT 2261
Oy 1380 GAGACCT 1439
Db 2262 GAGACCT 2321
Oy 1440 TGAAGCTCTGAG 1499
Db 2322 TGAAGCTCTGAG 1559
Oy 1500 TTTGCTTCTGAG 1559
Db 2382 TTTGCTTCTGAG 2441
Oy 1560 AAAAAAAAAA 1571
Db 2442 CACAAACATGAA 2453

RESULT 14
US-09-710-286-2850

Sequence 2850, Application US/09710286

GENERAL INFORMATION:

APPLICANT: Gearing, David P.

APPLICANT: McCarthy, Sean A.

APPLICANT: Holtzman, Douglas A.

TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

FILE REFERENCE: 1600, 2005-001

CURRENT APPLICATION NUMBER: US/09/710, 286

PRIOR FILING DATE: 2000-11-10

PRIOR APPLICATION NUMBER: 60/164, 255

NUMBER OF SEQ ID NOS: 4115

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2850
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
US-09-710-286-2850

Query Match 98.5%; Score 1548.8; DB 28; Length 2494;
Best Local Similarity 99.5%; Pred. No. 2.9e-298;
Matches 1564; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Oy 1 GCCCCACAGCCGCGCCCTCGAGCGCCCTGCGGCTACCGCGTAAAGAGCTGAAGCGAG 60
Db 882 GCCCCACAGCCGCGCCCTCGAGCGCCCTGCGGCTACCGCGTAAAGAGCTGAAGCGAG 941
Oy 61 AGCCGCGCCACAGCGGCGAGCG--CCCCACAGCGCCAGGAGACCCCTGCGAGCGCGCG 119
Db 942 AGCCGCGCCACAGCGGCGAGCGCCCGCCACAGCGGAGACCCCTGCGAGCGGAGCGCG 1001
Oy 120 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179
Db 1002 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1061
Oy 180 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
Db 1062 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
Oy 240 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
Db 1122 GGTGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
Oy 300 GAACACAGCGCGCGGAGCTGTGCGGCTGCGAGAGCTGTGCGGCTGCGAGCGCTGTGCT 359
Db 1182 GAACACAGCGCGCGGAGCTGTGCGGCTGCGAGAGCTGTGCGGCTGCGAGCGCTGTGCT 1241
Oy 360 CATGCTGAG 419
Db 1242 CATGCTGAG 1301
Oy 420 GAGAGCGGCGATCCAG 479
Db 1302 GAGAGCGGCGATCCAG 1361
Oy 480 AGTTCCTTGAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
Db 1362 AGTTCCTTGAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
Oy 540 CTGACAGCTATGCTGTGCGCGCGGCTGTGCTGATGATGATGATGATGATGATGATGAT 599
Db 1422 CTGACAGCTATGCTGTGCGCGCGGCTGTGCTGATGATGATGATGATGATGATGATGAT 1481
Oy 600 GGTTCGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 1482 GGTTCGCGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1541
Oy 660 GGAACCTAGAGAGTGAAGTATCGGCGCTGTGCGGAGATGCGCTTCACTGTGCGACCTT 719
Db 1542 GGAACCTAGAGAGTGAAGTATCGGCGCTGTGCGGAGATGCGCTTCACTGTGCGACCTT 1601
Oy 720 CCGCTGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
Db 1602 CCGCTGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1661
Oy 780 AGTGGTTTCAAG 839
Db 1662 AGTGGTTTCAAG 1721
Oy 840 CCTGTGCTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 899
Db 1722 CCTGTGCTCCAGACACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1781
Oy 900 TGTGCTAGTCTGAGACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 959
Db 1782 TGTGCTAGTCTGAGACCTGCTCCCGGTTGGTCCAGACAGGCGCGGTAGATGCACA 1841


```

Db 1962 CTACTGTGATATGTCGCCGCTGTCGCCCTTCGCTTGAGACCCAGATGGAAAGCTGT 2021
OY 1140 GTTTGCGAGTGATGGGGAATTTGATGTTTACGAGGCCGTGCAGGGCCAGTGCACCCAAA 1199
    |||
Db 2022 GTTTGCGAGTGATGGGGAATTTGATGTTTACGAGGCCGTGCAGGGCCAGTGCACCCAAA 2081
    |||
OY 1200 CTACTGTGATGTCAGCGGTTTGCTGAGACCCGCCAGCTGGAAGCCGCCAGCAGAT 1259
    |||
Db 2082 CTACTGTGATGTCAGCGGTTTGCTGAGACCCGCCAGCTGGAAGCCGCCAGCAGAT 2141
    |||
OY 1260 GCCACGCCAGAGAGCCCTTATGACCCCTGGGGCGGCTGTGCTTATGTCCTACTTGC 1319
    |||
Db 2142 GCCACGCCAGAGAGCCCTTATGACCCCTGGGGCGGCTGTGCTTATGTCCTACTTGC 2201
    |||
OY 1320 AGGACCCCTTCCTCCTTCCTAGGGCTGCAGGGCCCTGTCCAGACTCTGTGGGGGTGGAG 1379
    |||
Db 2202 AGGACCCCTTCCTCCTTCCTAGGGCTGCAGGGCCCTGTCCAGACTCTGTGGGGGTGGAG 2261
    |||
OY 1380 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439
    |||
Db 2262 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2321
    |||
OY 1440 TGAAGTCTCTGCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
    |||
Db 2322 TGAAGTCTCTGCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2381
    |||
OY 1500 TTTGTTCTGAGACCCGCCAGCCAGCAACCAATCAATTAAGTGACATTCGCCAGCTGA 1559
    |||
Db 2382 TTTGTTCTGAGACCCGCCAGCCAGCAACCAATCAATTAAGTGACATTCGCCAGCTGA 2441
    |||
OY 1560 AAAAAAAAAA 1571
    |||
Db 2442 CACAAACATCAA 2453
    |||

```

Search completed: May 17, 2003, 15:40:37
Job time : 3804 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 11:14:26 ; Search time 833 Seconds
(without alignments)
9771.695 Million cell updates/sec

Title: US-09-937-060A-19
Perfect score: 1573
Sequence: 1 gccccacagcgccgctctgcg.....gctgtaaaaaaaaaaaaaa 1573

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 6404235 seqs, 2587356060 residues

Total number of hits satisfying chosen parameters: 12808470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2-6/ptodata/2/pna/PCR_NEW_COMB.seq:*
2: /cgn2-6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2-6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2-6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2-6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2-6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2-6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2-6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgn2-6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgn2-6/ptodata/2/pna/US60_NEW_COMB.seq:*
11: /cgn2-6/ptodata/2/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1550.8	98.6	1783	US-09-949-016-1155	Sequence 1155, Ap
2	1549.8	98.5	1799	PCT-US03-02588-25	Sequence 25, Appl
3	1549.8	98.5	1799	US-10-354-358-25	Sequence 25, Appl
4	1459.8	92.8	1764	US-09-949-016-3940	Sequence 3940, Ap
5	1430.2	90.9	1562	US-10-264-237-1180	Sequence 1180, Ap
6	1414	89.9	1806	US-10-170-235-1781	Sequence 1781, Ap
7	1412	89.8	1806	US-60-453-135-5904	Sequence 5904, Ap
8	1412	89.8	1806	US-60-453-050-5904	Sequence 5904, Ap
9	1150.4	73.1	1152	US-09-949-016-12897	Sequence 22, Appl
10	1056	67.1	7198	US-09-949-016-15682	Sequence 12897, A
11	1056	67.1	7198	US-09-949-016-15682	Sequence 12897, A
12	1056	67.1	7198	US-09-949-016-15682	Sequence 12897, A
13	856.2	54.4	1927	US-10-144-771-2596	Sequence 2596, App
14	824.8	52.4	2648	US-10-338-044-2517	Sequence 2517, App
15	800.6	38.2	601	US-09-949-016-40620	Sequence 40620, A
16	600.6	38.2	601	US-09-949-016-40620	Sequence 40620, A
17	549.2	34.9	566	US-09-534-850-1813	Sequence 1813, Ap
18	439	27.9	523	US-09-534-850-1797	Sequence 1797, Ap
19	391.4	24.9	404	US-09-534-850-1812	Sequence 1812, Ap
20	328	20.9	376	US-09-912-293-138510	Sequence 138510, A
21	323.2	20.5	389	US-09-912-293-138513	Sequence 138513, A
22	305	19.4	429	US-09-912-293-138514	Sequence 138514, A

23	303.4	19.3	305	6	US-09-534-850-19644	Sequence 19644, A
24	291	18.5	480	6	US-09-912-293-214689	Sequence 214689, A
25	287	18.2	288	6	US-09-534-850-19657	Sequence 19657, A
26	273	17.4	275	6	US-09-534-850-19648	Sequence 19648, A
27	270	17.2	270	6	US-09-534-850-19656	Sequence 19656, A
28	265.8	16.9	295	6	US-09-534-850-19650	Sequence 19650, A
29	264	16.8	264	6	US-09-534-850-19659	Sequence 19659, A
30	262.8	16.7	1987	8	US-10-170-235-20325	Sequence 20325, A
31	262.8	16.7	1987	11	US-60-453-135-144	Sequence 135-144, App
32	262.8	16.7	1987	11	US-60-453-135-144	Sequence 135-144, App
33	262.8	16.7	1987	11	US-60-453-135-144	Sequence 135-144, App
34	262.8	16.7	2030	8	US-10-170-235-20475	Sequence 20475, A
35	262.8	16.7	2030	11	US-60-452-680-196	Sequence 196, App
36	262.8	16.7	2030	11	US-60-453-135-146	Sequence 146, App
37	262.8	16.7	2030	11	US-60-453-135-146	Sequence 146, App
38	262.8	16.7	2191	8	US-10-170-235-20327	Sequence 20327, A
39	262.8	16.7	2191	11	US-60-452-680-193	Sequence 193, App
40	262.8	16.7	2191	11	US-60-453-135-143	Sequence 143, App
41	262.8	16.7	2191	11	US-60-453-135-143	Sequence 143, App
42	262.8	16.7	2380	1	PCT-US03-02588-77	Sequence 77, Appl
43	262.8	16.7	2380	9	US-10-354-358-77	Sequence 77, Appl
44	259	16.5	260	6	US-09-534-850-19633	Sequence 19633, A
45	257	16.3	257	6	US-09-534-850-19661	Sequence 19661, A

ALIGNMENTS

```

RESULT 1
US-09-949-016-1155
; Sequence 1155, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1155

Query Match
Best Local Similarity 98.6%; Score 1550.8; DB 7; Length 1783;
Matches 1566; Conservative 0; Mismatches 7; Indels 1; Gaps 1:
1 GCCCCACAGCCGCGCTGCGAGCCCGCTGGCGACACCGATAGGAGCAGG 60
|||||
207 GCCCCACAGCCGCGCTGCGAGCCCGCTGGCGACACCGATAGGAGCAGG 266
|||||
61 ACCCGCGCGCGAGCGAGCG-CCCGACAGCGCGAGGAGCCCTGGCGAGCGCG 119
|||||
267 ACCCGCGCGCGAGCGAGCG-CCCGACAGCGCGAGGAGCCCTGGCGAGCGCG 326
|||||
120 GGTGCGAGTTATGATTCACAGCGGGCGCGCGCGCTGCTCCGCGCT 179
|||||
327 GGTGCGAGTTATGATTCACAGCGGGCGCGCGCGCGCTGCTCCGCGCT 386
|||||
180 GCTGCTGCTGCTGACACCGCGCGCGCGCAAGGCGCTTTCGCGAGTCA 239
|||||
387 GCTGCTGCTGCTGACACCGCGCGCGCGCAAGGCGCTTTCGCGAGTCA 446
|||||
240 CCGCAGCGCCCTTTGGCGCTGAGCGCTGAATCTCTTCACGCTGATGCTCAGGAGCGG 299

```

Db	447	CGTGCAGCCCTCTTTTGGCTGAGCCCTGAAATCTCTTCACTGAGTATGCTCACTGAGCGGG	506
Qy	300	GAACCAAGCGGGGAGGCTGATGCGGCTGAGAGAGCTGGGGCCGCTGGACCTCGTGGTG	359
Db	507	GAACCAAGCGGGGAGGCTGATGCGGCTGAGAGAGCTGGGGCCGCTGGACCTCGTGGTG	566
Qy	360	CATGTCAGAGACGGGCTGATGACAGAGGTGGTGAACGGGCTCAATGAGAGGGCTGACGTG	419
Db	567	CATGTCAGAGACGGGCTGATGACAGAGGTGGTGAACGGGCTCAATGAGAGGGCTGACGTG	626
Qy	420	GAGAGCCGGCATTCAGAAAGCCCGTGTAACCTCCACAGAGCTGTGGCAAGCGCTGGCG	479
Db	627	GAGAGCCGGCATTCAGAAAGCCCGTGTAACCTCCACAGAGCTGTGGCAAGCGCTGGCG	686
Qy	480	AGCTTCCTGAACCAATTATGTGGCTATAGACAGGTCAACCAATGAAGACCTCTGACCAA	539
Db	687	AGCTTCCTGAACCAATTATGTGGCTATAGAGCAAGGTCAACCAATGAAGACCTCTGACCAA	746
Qy	540	CTGCACGCTATTGCTGTGCGCGCGGCTGCTGTACCAATGAACCTGCTCTGCACAC	599
Db	747	CTGCACGCTATTGCTGTGCGCGCGGCTGCTGTACCAATGAACCTGCTGCTGCACAC	806
Qy	600	GAGCTTGGGGGCTGCACCTCTCTCTCTGTGCTCAACCTGGCGCGGGGCTTCAATTGCTATGT	659
Db	807	GAGCTTGGGGGCTGCACCTCTCTCTCTGTGCTCAACCTGGCGCGGGGCTTCAATTGCTATGT	866
Qy	660	GGAACCTAGAGAGTGAAGTATCGGCTCTGGGAGAGATGGGCTTCACTGTGGGACACTT	719
Db	867	GGAACCTAGAGAGTGAAGTATCGGCTCTGGGAGAGATGGGCTTCACTGTGGGACACTT	926
Qy	720	CCAGGCTCTGGAGAGCCCTGAGCACCTACCGGAGCGACCTGAGGCTACCTCCCTGTAGAAG	779
Db	927	CCAGGCTCTGGAGAGCCCTGAGCACCTACCGGAGCGACCTGAGGCTACCTCCCTGTAGAAG	986
Qy	780	AGTGGGTTCCAGAGACCTGCTGCTCCCGTTGGTGCACAGGAGGCGCGGTAAATGACACA	839
Db	987	AGTGGGTTCCAGAGACCTGCTGCTCCCGTTGGTGCACAGGAGGCGCGGTAAATGACACA	1046
Qy	840	CGTTTGGGCACTGGAGAGAGCCAGTGGCTCTCTCACTGAGAGTGGTGGCCGACAGAGACTT	899
Db	1047	CGTTTGGGCACTGGAGAGAGCCAGTGGCTCTCTCACTGAGAGTGGTGGCCGACAGAGACTT	1106
Qy	900	TGAGTCATGTCGGACATGCGACGTGCACTGCGACCGGAGCAATGAGATGTTTCTGACCACT	959
Db	1107	TGAGTCATGTCGGACATGCGACGTGCACTGCGACCGGAGCAATGAGATGTTTCTGACCACT	1166
Qy	960	GGGGCGGCTGTGCAGCTGGCGTCAATCATCTCTTCTTCACTGCGGGCGGAGTGTCTCTCG	1015
Db	1167	GGGGCGGCTGTGCAGCTGGCGTCAATCATCTCTTCTTCACTGCGGGCGGAGTGTCTCTCG	1226
Qy	1020	CATGCTGCTGCAGCTCTTCTCTGAGGCATGAGAGGAGGAGGAGGCAATGAGATGAATGCC	1075
Db	1227	CATGCTGCTGCAGCTCTTCTCTGAGGCATGAGAGGAGGAGGAGGCAATGAGATGAATGCC	1286
Qy	1080	CTACTCTTGATATATGTGCCCGCTGCTGCGCTCTCGCTTGAAGCCCAAGAGATGGGAAAGTGT	1133
Db	1287	CTACTCTTGATATATGTGCCCGCTGCTGCGCTCTCGCTTGAAGCCCAAGAGATGGGAAAGTGT	1346
Qy	1140	GTTTTCAGATGAGTGGGAAATTGATGCTTACGAGCGCGTGCACGGGCGGACAGTGCACCAA	1199
Db	1347	GTTTTCAGATGAGTGGGAAATTGATGCTTACGAGCGCGTGCACGGGCGGACAGTGCACCAA	1406
Qy	1200	CTACTCTTGATATATGTGCCCGCTGCTGAGGAGCCCGCGGCAAGTGGAAAGGCCAGAGAT	1255
Db	1407	CTACTCTTGATATATGTGCCCGCTGCTGAGGAGCCCGCGGCAAGTGGAAAGGCCAGAGAT	1466
Qy	1260	GCAACGCGCAGAAAGCCCTTATGACCCCTGGGCGACAGGTGTGCTTAAAGTCTACTCTGC	1313
Db	1467	GCAACGCGCAGAAAGCCCTTATGACCCCTGGGCGACAGGTGTGCTTAAAGTCTACTCTGC	1526
Qy	1320	AGGACCTTCTCTCTTCCCTAGGGCTGCAGGGCTGTCCACAAGCTCTGTGGGGGTGAG	1375

Db	1527	AGGAGCCCTTCCTCCCTCCCTCCAGGGCCGACAGGACCTGTCACAGTCTCTGAGGGGGGAG	1586
QY	1380	GAGACTCCTCTGGAAAGAGGCGAGAAAGTGGAGGCTATGCTTTGGGGGACAGGCCAGAA	1439
Db	1587	GAGACTCCTCTGGAAAGAGGCGAGAAAGTGGAGGCTATGCTTTGGGGGACAGGCCAGAA	1646
QY	1440	TGAGAGTCGCGGGTCAGAGCCGACGACCTGCGCTGGGGCCACATCCCTATGTAAGGCTCTTAG	1499
Db	1647	TGAGAGTCGCGGGTCAGAGCCGACGACCTGCGCTGGGGCCACATCCCTATGTAAGGCTCTTAG	1706
QY	1500	TTTTGCTGTGAGAGCCCCCAGCCCGCAGAACCAATTCGAATTAAGAGCATTTCCAGGCTGGA	1559
Db	1707	TTTTGCTGTGAGAGCCCCCAGCCCGCAGAACCAATTCGAATTAAGAGCATTTCCAGGCTGGA	1766
QY	1560	AAAAAAAAAAAAAA 1573	
Db	1767	AAAAAAAAAAAAAA 1780	

```

RESULT 2
PCT-US03-02588-25
Sequence 25, Application PC/TUS0302588
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2088, 2427, 2702, 5891, 6428,
TITLE OF INVENTION: 7181, 7650, 25641, 69583, 149863, 1866927, 1682, 17637, 2325,
TITLE OF INVENTION: 7183, 14171, 103356, 16580, 1480, 18984, 2088, 32937, 24580,
TITLE OF INVENTION: 8952, 8989, 1642, 68268, 10287, 1584, 9225, 14124, 44680,
TITLE OF INVENTION: 8950, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943
TITLE OF INVENTION: 16334, 26888, 9011, 14031, 6178, 212225, 1420, 32236, 2095
TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 65428 MOLECULE
FILE REFERENCE: MP102.0020
CURRENT APPLICATION NUMBER: PCT/US03/02588
CURRENT FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See file wrapper or PALM.
Number of SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 1799
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (359)...(1513)
PCT-US03-02588-25

```

Query Match 98.5%; Score 1549.8; DB 1; Length 1799;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1562; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

1 GCGCCACACCGCCCTGAGAGCCCGCTGGGAGACACCGATTAAGAGCTGAAGCAG 60
 229 GCGCCACACCGCCCTGAGAGCCCGCTGGGAGACACCGATTAAGAGCTGAAGCAG 288
 61 ACCCGCCGCGAGGAGAGG-CCCGACAGCGCAGGAGACCCCTGAGCAGGAGCGG 119
 289 ACCCGCCGCGAGGAGAGGCCCCCAGACAGCGCAGGAGACCCCTGAGCAGGAGCGG 348
 120 GGTGAGGTTATGATGACAGCGGCGCGCCCGCGGCGGTCTCCCGCGCCCTGCCGCT 179
 349 GGTGAGGTTATGATGACAGCGGCGCGCCCGCGGCGGTCTCCCGCGCCCTGCCGCT 408
 180 GGTGAGGTTATGATGACAGCGGCGCGCGCGCGGCGGTCTCCCGCGCCCTGCCGCT 239
 409 GGTGAGGTTATGATGACAGCGGCGCGCGCGCGGCGGTCTCCCGCGCCCTGCCGCT 468
 240 GGTGAGGTTATGATGACAGCGGCGCGCGCGCGGCGGTCTCCCGCGCCCTGCCGCT 299
 469 GGTGAGGTTATGATGACAGCGGCGCGCGCGCGGCGGTCTCCCGCGCCCTGCCGCT 528
 300 GAACACGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
 529 GAACACGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
 360 CATGTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
 589 CATGTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 648
 420 GAGAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
 649 GAGAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 708
 480 AGCTTCCTTGAACCAATATGCTGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAG 539
 709 AGCTTCCTTGAACCAATATGCTGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAG 768
 540 CTGACGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
 769 CTGACGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 828
 600 GGTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
 829 GGTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 888
 660 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719
 889 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
 720 CCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779
 949 CCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
 780 AGTGGTTCAAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 839
 1009 AGTGGTTCAAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1068
 840 CCTTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
 1069 CCTTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1128
 900 TGTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 959
 1129 TGTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1188
 960 GAGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1019
 1189 GAGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1248

1020 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1079
 1249 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1308
 1080 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1139
 1309 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1368
 1140 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1199
 1369 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1428
 1200 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1259
 1429 CATGCTGCTGCGGCTCTCTCTGCGCATGAGAGAGGCGAGGAGGATGAGATGATGAGCC 1488
 1260 GCGAGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1319
 1489 GCGAGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1548
 1320 AGGAGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1379
 1549 AGGAGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1608
 1380 GAGACTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1439
 1609 GAGACTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
 1440 TGAAGTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1499
 1669 TGAAGTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1728
 1500 TTTGTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1559
 1729 TTTGTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1788
 1560 AAAAAAAAAA 1570
 1789 AAAAAAAAAA 1799

RESULT 3
 US-10-354-25
 ; Sequence 25, Application US/10354358
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc
 ; APPLICANT: Hunter, John Joseph
 ; APPLICANT: Macbeth, Kyle J.
 ; APPLICANT: Tsai, Fong-Ying
 ; APPLICANT: Lesoon, Andrea
 ; APPLICANT: Lightcap, Eric S.
 ; APPLICANT: Williamson, Mark
 ; APPLICANT: Rudolph-Owen, Laura A.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
 ; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 16842, 17667, 9235,
 ; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
 ; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
 ; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10460, 20893, 33230, 1586, 9943,
 ; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
 ; FILE REFERENCE: MP102-020P1R0MNM1M
 ; CURRENT APPLICATION NUMBER: US/10354, 358
 ; CURRENT FILING DATE: 2003-01-30
 ; PRIOR APPLICATION NUMBER: US 60/353, 600
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: US 60/364, 517
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: US 60/371, 075
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: US 60/371, 507
 ; PRIOR FILING DATE: 2002-04-10
 ; PRIOR APPLICATION NUMBER: US 60/372, 984


```

: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 3940
: LENGTH: 1764
: TYPE: DNA
: ORGANISM: Human
: OS-03-949-016-3940

```

Query Match	92.88;	Score 1459.8;	DB 7;	Length 1764;
Best Local Similarity	97.38;	Pred. No. 8e-312;		
Matches 1517; Conservative	0;	Mismatches 37;	Indels 5;	Gaps 3

QY	1	GCCCACAGCGGCGCTGGAGACGCCCGCTGGGACACACGATTAAAGAGCTCAAGCGAG	60
Db	207	GCCCACAGCGGCGCTGGAGACGCCCGCTGGGAGACACCGATTAAAGAGCTCAAGCGAG	266
QY	61	AGCCGCGCCACAGGGGAGCG - CCGCACAGCGCCAGGAGCCCTCTGACAGCGGGACCGCG	119
Db	267	AGCCGCGCCACAGGGGAGCGCCCGCCACAGCGCCAGGAGCCCTCTGGACAGCGGGACCGCG	326
QY	120	GGTAGAGATTATGGATTCACAGGGGCGGCCCGCGGGCGTGTCTCCGCGGCGCTCGCCGCT	179
Db	327	GGTAGAGATTATGGATTCACAGAGGGGCGGCCCGCGGGCGTAGCTCCGCGGGCGCTTGCTTA	386
QY	180	GCTGTGCTCTCTGACACCGCGGGCGGGACAGGAGCGCTTGACAGCTTTCGGAGTCA	239
Db	387	TGAATGACTGTGACAC - GCGAGAGGCAAGGAGCAAGGC - TTCAAGCTTTCGGAGTCA	442
QY	240	CGTGCAGCCCTTTTGGCTGAGGCTGTAAATCTCTTCACTGATATGCTACAGAGCGCG	299
Db	443	CGTGCAGCCCTTTTGGCTGAGGCTGTAAATCTCTTCACTGATATGCTACAGAGCGCGCG	502
QY	300	GAACACAGCGCGGGAGCTGTGTGCGGAGAGCTGGGCGCGCTGGAGCGCTGTGTGT	359
Db	503	GAACACAGCGCGGGAGCTGTGTGCGGAGAGCTGGGCGCGCTGGAGAGCTGTGTGTGT	562
QY	360	CATGTCTGAGAGAGCGGCTGTATGACAGAGGTGTGAACGGGCTCATAGAGAGCGGCTGAGCTG	419
Db	563	CATGTCTGAGAGAGCGGCTGTATGACAGAGGTGTGAACGGGCTCATAGAGAGCGGCTGAGCTG	622
QY	420	GGAACACGGCATTCAGAAAGCCCTGTGTATGCTCCAGACGGCTCTGGACAGCGCTGGC	479
Db	623	GGAACACGGCATTCAGAAAGCCCTGTGTATGCTCCAGAGCGGCTCTGGACAGCGCTGGC	682
QY	480	AGCTTCCCTTGAACCATTAATGCTGGATATGAGAGCGGTGACCATAGAAAGCTCTGACGAA	539
Db	683	AGCTTCCCTTGAACCATTAATGCTGGATATGAGAGCGGTGACCATTAAGAGACTCTCTACGAA	742
QY	540	CTGAGAGCTATTGCTGTGCCGCGCGGCTGCTGTACCCATGAACCTGTGTCTCTGACAGC	599
Db	743	CTGAGAGCTATTGCTGTGCCGCGCGGCTGCTGTACCCATGAACCTGTGTCTCTGACAGC	802
QY	600	GGCTTGGGGGCTGTGGGCTCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	659
Db	803	GGCTTGGGGGCTGTGGGCTCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	862
QY	660	GGACCTTAGAGAGTGAAGAAGATATGGGGGCTGGGGGAGATAGCGTTCACTGTGGGAGACCT	719
Db	863	GGACCTTAGAGAGTGAAGAAGATATGGGGGCTGGGGGAGATAGCGTTCACTGTGGGAGACCT	922
QY	720	CTGTGCTGTGGACCGCTGGGCACTAACCGGGGCGAGTATGGCTATCTCCCTGTAGAGAG	779
Db	923	CTGTGCTGTGGACCGCTGGGCACTAACCGGGGCGAGTATGGCTATCTCCCTGTAGAGAG	982
QY	780	AGTAGGTTTCAAGACACTCTCTCCCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	839
Db	983	AGTAGGTTTCAAGACACTCTCTCCCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1042

QY	840	CGTTGTGGCACTGGAGAGGACGAGTGGCCCTTCACCTGACAGTGGTGGCCGACGAGAGACTT	899
Db	1043	CGTTGTGGCACTGGAGAGGACGAGTGGCCCTTCACCTGACAGTGGTGGCCGACGAGAGACTT	1107
QY	900	TGTGTATGCTCTGGACACGTGCTGACACTGACACGTGGGACAGTGAAGATTTTGTCTGACCCAT	959
Db	1103	TGTGTATGCTCTGGACACGTGCTGACACTGACACGTGGGACAGTGAAGATTTTGTCTGACCCAT	1166
QY	960	GAGCGCGCTGTGACAGCTGGCGTCATGCATCTGTTCTTACGTGCGGGGGGAGTGTCTTCGTGC	1019
Db	1163	GAGCGCGCTGTGACAGCTGGCGTCATGCATCTGTTCTTACGTGCGGGGGGAGTGTCTTCGTGC	1222
QY	1020	CATGTGCTGCGGCTCTTCTCCGGGCAATGAGGAAGAGGACAGGCATATGAGATTAATGAATGGCC	1079
Db	1223	CATGTGCTGCGGCTCTTCTCCGGGCAATGAGGAAGAGGACAGGCATATGAGATTAATGAATGGCC	1288
QY	1080	CTACTTGGTATATGTGCCCGTGCTGCTTCGCTTCGCTTGGAGGCCCAAGCATGGGAAAGTGCT	1139
Db	1283	CTACTTGGTATATGTGCCCGTGCTGCTTCGCTTCGCTTGGAGGCCCAAGCATGGGAAAGTGCT	1344
QY	1140	GTTTGTGACATGATGGGGAATTGATGCTTACGAGAGCCGTGACAGGCGCAGGTGACCCCAAA	1199
Db	1343	GTTTGTGACATGATGGGGAATTGATGCTTACGAGAGGCGGTGACAGGCGCAGGTGACCCCAAA	1402
QY	1200	CTACTTGTGATGTGCAGAGGTTTCCGTGGAGGCCCGCGCCAGCTGTGAAACCCCGACAGCAT	1255
Db	1403	CTACTTGTGATGTGCAGAGGTTTCCGTGGAGGCCCGCGCCAGCTGTGAAACCCCGACAGCAT	1466
QY	1260	GCCACCGCCAGAAAGAGCCCTTATGACCCCTGGGCGGCGGTGCTTGTACTATTCGC	1319
Db	1463	GCCACCGCCAGAAAGAGCCCTTATGACCCCTGGGCGGCGGTGCTTGTACTATTCGC	1522
QY	1320	AGGACCGCTTCTCTTCTTCCCTTAAGGGCTCCAGGGCTGTCCACAGCTCTGTGGGGGGTGGAG	1379
Db	1523	AGGACCGCTTCTCTCTTCCCTTAAGGGCTCCAGGGCTGTCCACAGCTCTGTGGGGGGTGGAG	1583
QY	1380	GAGACTCTCTGAGAAAGGAGGAGAAAGCTGAGAGCTATGCTTTGGGGGACAGAGCCAGAA	1433
Db	1583	GAGACTCTCTGAGAAAGGAGGAGGAGAAAGCTGAGAGCTATGCTTTGGGGGACAGAGCCAGAA	1644
QY	1440	TGAAGTCTCTGGGTGAGAGCCAGCTGGTGGGCCACAGCTGCCTATGTAAGGCTTTCTAG	1499
Db	1643	TGAAGTCTCTGGGTGAGAGCCAGCTGGTGGGCCACAGCTGCCTATGTAAGGCTTTCTAG	1702
QY	1500	TTTGTTCGAAACCCCAACCCCAAGAACCAATTCAAATTAATGACATTTCCAGCGCTG	1558
Db	1703	TTTGTTCGAAACCCCAACCCCAAGAACCAATTCAAATTAATGACATTTCCAGCGCTG	1761
RESULT 5			
US-10-264-237-1180/c			
; Sequence 1180, Application US/10264237			
; GENERAL INFORMATION:			
; APPLICANT: Birse et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies			
; FILE REFERENCE: PA131P1			
; CURRENT APPLICATION NUMBER: 2002-10-04			
; PRIOR APPLICATION NUMBER: PCT/US01/16450			
; PRIOR FILING DATE: 2001-05-18			
; PRIOR APPLICATION NUMBER: US 60/205,515			
; PRIOR FILING DATE: 2000-05-19			
; NUMBER OF SEQ. ID NOS: 2876			
; SOFTWARE: PatentIn Ver. 3.1			
; SEQ ID NO 1180			
; LENGTH: 1562			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (7)..(7)			
; OTHER INFORMATION: n equals a..t..g, or c			

Db 446 --CGCGAAGGCAAGGGGCAAGGC--TTACAGCTCTTCGGAGTCACTGTGCAAGCCCTTTTGG 501
OY CTGAGAGCTGAAATCTCTCTTACAGCTGATGCTCACTGAGCGGCGGAACCAAGCGCGGAGG 316
Db 502 CTGAGAGCTGAAATCTCTCTTACAGCTGATGCTCACTGAGCGGCGGAACCAAGCGCGGAG 501
OY 317 TGGTGGGCTGAGAGAGCTGGGCGGCTGGAGACCTGTGGTGCATGTCTGAGAGCGGCG 376
Db 562 GGTGGGATCGGAGAGCTGGGCGGCTGAGAGAGCTGTGGTGCATGTCTGAGAGCGGCG 621
OY 377 TGATGACAGAGGTGTGAAGCGGCTCATGAGAGGCGCTGACTGGAGACCGCATCCAGA 436
Db 622 TGATGACAGAGGTGTGAAGCGGCTCATGAGAGGCGCTGACTGGAGACCGCATCCAGA 691
OY 437 AGCCCTGTGTAGCCCTCCAGAGAGGCTGTGCAAGCGGCTGGAGAGCTTCCTTGAACCAT 496
Db 682 AGCCCTGTGTAGCCCTCCAGAGAGGCTGTGCAAGCGGCTGGAGAGCTTCCTTGAACCAT 741
OY 497 ATGCTGGCTATGAGAGAGGTGACCAATGAAGACCTCTGACCACTGACAGCTATTGCTGT 556
Db 742 ATGCTGGCTATGAGAGAGGTGACCAATGAAGACCTCTGACCACTGACAGCTATTGCTGT 801
OY 557 GCGCGCGGCTGCTGTACACCATGAACCTGCTGTCTGACACAGCGCTTGGGCGCTCGCC 616
Db 802 GCGCGCGGCTGCTGTACACCATGAACCTGCTGTCTGACACAGCGCTTGGGCGCTCGCC 861
OY 617 TCTTCTGTGTGCTCAGCCTTGCGCTTGGGCGCTTCACTGTGATGATGAGAGTGA 676
Db 862 TCTTCTGTGTGCTCAGCCTTGCGCTTGGGCGCTTCACTGTGATGATGAGAGTGA 921
OY 677 AGTATGCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGTGGACGCC 736
Db 922 AGTATGCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGTGGACGCC 981
OY 737 TGGGCAACCAACCGCGGCGGCTGAGCTTACCTTCCCTTAGAGAGAGTGGTTCACACAC 796
Db 982 TGGGCAACCAACCGCGGCGGCTGAGCTTACCTTCCCTTAGAGAGAGTGGTTCACACAC 1041
OY 797 CTGCGCTCCCGCTGTGTGTGCTGACAGAGCGCGGCTAGATGACACCTTGTGCTGAGG 856
Db 1042 CTGCGCTCCCGCTGTGTGTGCTGACAGAGCGCGGCTAGATGACACCTTGTGCTGAGG 1101
OY 857 AGCAGCTGCCCTCTCACTGAGACAGTGTGTGCGCCAGAGAGCATTTGTGTAGTCTGGAC 916
Db 1102 AGCAGCTGCCCTCTCACTGAGACAGTGTGTGCGCCAGAGAGCATTTGTGTAGTCTGGAC 1161
OY 917 TGTGTGACACCTGAGAGAGT 976
Db 1162 TGTGTGACACCTGAGAGAGT 1221
OY 977 GCGTCATGATCTGTCTCACTGAGTGTGGGCGGAGTGTCTGTGCTGATGCTGTGCGCTCT 1036
Db 1222 GCGTCATGATCTGTCTCACTGAGTGTGGGCGGAGTGTCTGTGCTGATGCTGTGCGCTCT 1281
OY 1037 TCTGTGCACTGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGATG 1096
Db 1282 TCTGTGCACTGAGAGAGGAGCATATGAGATGATGATGATGATGATGATGATGATGATG 1341
OY 1097 CCGTGTGCTGCTTCCGCTGTGAGACCAAGATGTGAAGGTGTGTGTGTGTGTGTGTGTGT 1156
Db 1342 CCGTGTGCTGCTTCCGCTGTGAGACCAAGATGTGAAGGTGTGTGTGTGTGTGTGTGTGT 1401
OY 1157 AATTGATGTTAGAGAGCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1216
Db 1402 AATTGATGTTAGAGAGCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1461
OY 1217 GCGGTTGCTGAGAGCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1276
Db 1462 GCGGTTGCTGAGAGCGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1521
OY 1277 CATTATGACCCCTGAGGCGGCTGTGCTTACTGCTTACTGCTGAGAGAGCTTCTCTTCT 1336
Db 1522 CATTATGACCCCTGAGGCGGCTGTGCTTACTGCTTACTGCTGAGAGAGCTTCTCTTCT 1581

OY 1337 CATTAGGCTGACAGAGGCTGTGACAGCTCTGCTGGGCTGAGAGAGCTCTGTGAGAA 1396
Db 1582 CATTAGGCTGACAGAGGCTGTGACAGCTCTGCTGGGCTGAGAGAGCTCTGTGAGAA 1641
OY 1397 GAGTGAAGAGTGTGAGAGCTATGCTTTGGGGGACAGAGCCAGATGAAGTCTGGTCA 1456
Db 1642 GAGTGAAGAGTGTGAGAGCTATGCTTTGGGGGACAGAGCCAGATGAAGTCTGGTCA 1701
OY 1457 AGCCAGAGTGTGGGCGGCTGAGAGGCTGAGAGGCTTGTAGTGTGTGTGTGTGTGTGT 1516
Db 1702 AGCCAGAGTGTGGGCGGCTGAGAGGCTGAGAGGCTTGTAGTGTGTGTGTGTGTGTGT 1761
OY 1517 ACCCCAGAGAGCAATTCGAATGAAGTGAACATTCAGAGCTG 1558
Db 1762 ACCCCAGAGAGCAATTCGAATGAAGTGAACATTCAGAGCTG 1803

RESULT 7

US-60-453-135-5904

Sequence 5904, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOBBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01456

CURRENT APPLICATION NUMBER: US/60/453,135

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5904

LENGTH: 1806

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-5904

Query Match 89.8%; Score 1412; DB 11; Length 1806;

Best Local Similarity 95.1%; Pred. No. 2, 8e-301;

Matches 1523; Conservative 5; Mismatches 25; Indels 49; Gaps 5;

OY 1 GCCCCACAGCGGCGCTGCGAGCGCCGCTGGGACAGCAGGATGAAGAGTGAAGCAGG 60
Db 207 GCCCCACAGCGGCGCTGCGAGCGCCGCTGGGACAGCAGGATGAAGAGTGAAGCAGG 266
OY 61 AGCCCGCGCCAGCGGAGAGG-CCCCACAGCGCCAGGAGCCCTTGGAGCGGAGCGCG 119
Db 267 AGCCCGCGCCAGCGGAGAGG-CCCCACAGCGCCAGGAGCCCTTGGAGCGGAGCGCG 326
OY 120 GGTGAGAGTGTATGATCCAG-----139
Db 327 GGTGAGAGTGTATGATCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 386
OY 140 --CGGCGGCGCCCGCGGCGGCTGCTCCCGCGGCGCTGCGCGGCTGCTGCTGCTGAGC 196
Db 387 AGCAGGCGGCGCCCGCGGCGGCTGCTCCCGCGGCGCTGCGCGGCTGCTGCTGAGC 445
OY 197 GCGGCGGCGGCGGAGGAGGCTTGTGAGCTTTCGGAATGACAGTGTGAGAGGCTTTGG 256
Db 446 --CGGCGGCGGAGGAGGAGG---TTGAGCTTTCGGAATGACAGTGTGAGAGGCTTTGG 501
OY 257 CTGAGGCTGAAATCTCTTACGCTGATGCTCACTGAGCGGCGGAGCAAGCGGCGGAGG 316
Db 502 CTGAGGCTGAAATCTCTTACGCTGATGCTCACTGAGCGGCGGAGCAAGCGGCGGAGG 561
OY 317 TGGTGGGCTGAGAGAGCTGGGCGGCTGAGAGGCTGTGATGATGCTTGAAGAGGCG 376
Db 562 GGTGGGATCGGAGAGCTGGGCGGCTGAGAGGCTGTGATGATGCTTGAAGAGGCG 621
OY 377 TGATGACAGAGGTGTGAAGCGGCTCATGAGAGGCGCTGACTGGAGAGCGCATCCAGA 436
Db 622 TGATGACAGAGGTGTGAAGCGGCTCATGAGAGGCGCTGACTGGAGAGCGCATCCAGA 681

QY	437	ACCACCTGTAGCTCCACAGAGGCTCTGGCAACCGCTGGACAGCTTCCTTGAACCAATT	498
Db	682	AACCCCTGTAGCTCCACAGAGGCTCTGGCAACCGCTGGACAGCTTCCTTGAACCAATT	741
QY	497	ATGCTGGCTATGACAGAGTCAACAATGAAACCTCTGACCACACTGACACTATTGGCT	556
Db	742	ATGCTGGCTATGACAGAGTCAACAATGAAACCTCTGACCACACTGACACTATTGGCT	801
QY	557	GGCGGGGCTGTGTACACCAATGAACCGTGTCTGTGACACGGCTTGGGGGCTGGCC	618
Db	802	GGCGGGGCTGTGTACACCAATGAACCGTGTCTGTGACACGGCTTGGGGGCTGGCC	861
QY	617	TCTTCTCTGTGTCTACGCTGGCTGGGGCTTCAATGTGTATGTGGACCTAGAGATGAGA	678
Db	862	TCTTCTCTGTGTCTACGCTGGCTGGGGCTTCAATGTGTATGTGGACCTAGAGATGAGA	921
QY	677	AGTATGGGCTGTGGGGAGATGCCCTTCACTGTGGGACACCTCTCTCGTGGACGCC	738
Db	922	AGTATGGGCTGTGGGGAGATGGCTTCACTGTGGGACACCTCTCTCGTGGAGGCC	981
QY	737	TGCGACACTACCGGGCGAGACTGTACCTTCCCTGTAGAGAGTGGGTCCAGAGAC	796
Db	982	TGCGACACTACCGGGCCAGACTGTACCTTCCCTGTAGAGAGTGGGTCCAGAGAC	1046
QY	797	CTGACCTCCCGTGTGTGTCTAGCAAGGAGCCCGGTATAGTCAACCTTGTGCCACTGGAG	856
Db	1042	CTGACCTCCCGTGTGTGTGTCTAGCAAGGAGCCCGGTATAGTCAACCTTGTGCCACTGGAG	1101
QY	857	AGCCAGTCCCTCTCACTGTGAGAGTGTGCCACAGACACTTGTCTGTAGTCTGTGGAC	918
Db	1102	AGCCAGTCCCTCTCACTGTGAGAGTGTGCCACAGAGACTTGTCTGTAGTCTGTGGAC	1168
QY	917	TGCTGACTGTCAACTGTGGGCACTGTGATGTTCCTCAACCATGGGCGGTGTGGAGCTG	978
Db	1162	TGCTGACTGTCAACTGTGGGCACTGTGATGTTCCTCAACCATGGGCGGTGTGGAGCTG	1222
QY	977	GGGTATGCACTGTCTCACTGTGGGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1039
Db	1222	GGGTATGCACTGTCTCACTGTGGGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1282
QY	1037	TGCTGTGCTATGAGAGGAGGCAAGCTATGAGATGTATGTGCTTGTGTGTGTGTGTGTGT	1098
Db	1282	TGCTGTGCTATGAGAGGAGGCAAGCTATGAGATGTATGTGCTTGTGTGTGTGTGTGTGT	1344
QY	1097	CCGT	1155
Db	1342	CCGT	1406
QY	1157	AATGTATGTTATGAGAGGCGCTGTGAGGCGCAGGTGACCCAAACTACTTGTGATGTCTA	1211
Db	1402	AATGTATGTTATGAGAGGCGCTGTGAGGCGCAGGTGACCCAAACTACTTGTGATGTCTA	1468
QY	1217	GGCTTTGGCTGTGAGACCCCGCCCACTGTGAAAGCCCGACAGATGTCCACGCGCAGAAAGGC	1277
Db	1462	GGCTTTGGCTGTGAGACCCCGCCCACTGTGAAAGCCCGACAGATGTCCACGCGCAGAAAGGC	1522
QY	1277	CCTTATGACCCCTGTGAGCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1333
Db	1522	CCTTATGACCCCTGTGAGCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1588
QY	1337	CCTTATGACCCCTGTGAGCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1393
Db	1582	CCTTATGACCCCTGTGAGCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1644
QY	1397	GGGTATGAGAGTGTGAGGCTATCTTGTGGGGGAGACAGCAGATGAAGTCTGTGGGTGAG	1458
Db	1642	GGGTATGAGAGTGTGAGGCTATCTTGTGGGGGAGACAGCAGATGAAGTCTGTGGGTGAG	1707
QY	1457	AGCCAGT	1511
Db	1702	AGCCAGT	1766
QY	1517	ACCCAGCAACCAATTCAAATTAATGATTCCTCCACCTTG	1558

OY 1163 TGGTTAGCGAGCCGTCAGAGGCGCAGTGTGACACCAACTACTTCTGGATGTCACGGCTT 1222
 DB 4624447 TGGTTAGCGAGCCGTCAGAGGCGCAGTGTGACACCAACTACTTCTGGATGTCACGGCTT 4624388
 OY 1223 GCGTGGAGCCCGCCGACCTGTGAGAGCCGAGGATGCCAGCCGAGAGAGCCCTAT 1282
 DB 4624387 GCGTGGAGCCCGCCGACCTGTGAGAGCCGAGGATGCCAGCCGAGAGAGCCCTAT 4624328
 OY 1283 GACCCCTGGGCGCCGCTGTGCTTACTGTCTACTTGTGAGAGCCCTTCTCTCTCTAGG 1342
 DB 4624327 GACCCCTGGGCGCCGCTGTGCTTACTGTCTACTTGTGAGAGCCCTTCTCTCTCTAGG 4624268
 OY 1343 GCTGCGAGGCTGTGTCACAGCTCTGTGGGGGTGAGAGATCTCTGTGAGAGAGGTGA 1402
 DB 4624267 GCTGCGAGGCTGTGTCACAGCTCTGTGGGGGTGAGAGATCTCTGTGAGAGAGGTGA 4624208
 OY 1403 GAAGGTGAGGCTATGCTTTGGGGGAGAGCCGAGAAATGATGCTGCTGAGAGCCCA 1462
 DB 4624207 GAAGGTGAGGCTATGCTTTGGGGGAGAGCCGAGAAATGATGCTGCTGAGAGCCCA 4624148
 OY 1463 GCTGCTGGGCGCCAGCTGTGCTATGTAAGGCTTCTAGTTGTTCTGAGAGCCGAGCCCA 1522
 DB 4624147 GCTGCTGGGCGCCAGCTGTGCTATGTAAGGCTTCTAGTTGTTCTGAGAGCCGAGCCCA 4624088
 OY 1523 CGAACCAATCCAAATTAAGTACATTCAGAGCCCTG 1558
 DB 4624087 CGAACCAATCCAAATTAAGTACATTCAGAGCCCTG 4624052

RESULT 13

US-10-144-771-2596
 : Sequence 2596 Application US/10144771
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig
 : TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
 : FILE REFERENCE: CL001321
 : CURRENT APPLICATION NUMBER: US/10/144,771
 : CURRENT FILING DATE: 2002-05-15
 : NUMBER OF SEQ ID NOS: 47235
 : SEQ ID NO 2596
 : LENGTH: 1927
 : TYPE: DNA
 : ORGANISM: HUMAN
 : US-10-144-771-2596

Query Match 54.4% Score 856.2; DB 9; Length 1927;
 Best Local Similarity 74.0%; Pred. No. 8,6e-179;
 Matches 1159; Conservative 0; Mismatches 388; Indels 20; Gaps .5;

OY 1 GCCCCACAGCGCGCCCTGTGGAGCGCCGCTGTGGAGCAGCAGATTAAGAGAGCTGAAGGCGAG 60
 DB 365 GCCCCACAGCGCGCCCTGTGGAGCGCCGCTGTGGAGCAGCAGATTAAGAGAGCTGAAGGCGAG 424
 OY 61 AGCGCGCGCCGACGAGAGCGCC---ACAGCGCCGAGGAGACCCCTGTGACAGCGAGAC 116
 DB 425 AGCGCGCGCTTACTGTAGAGAGCGCGGCGAGCAGCGGTGCGCCCTGTGACAGCGAGCC 484
 OY 117 GCGGCTGAGGCTTATGATCCAGCGCGCGCCCGCGGCGTGTCTCCGCGCGCTGCG 176
 DB 485 CCGGCACTGTGATGAGAACAGTAGAATGCTCTCGAGGACTGTCTCCGCGCGCTGAG 544
 OY 177 CGTGCTGCTGCTGCTGAGAACCGCGCGCGGCGGAGAGGAGAGGAGCTTGTGACGCTTCCGAG 236
 DB 545 AGTCTGCTGCTGCTGAGAACCGCGCGGAGGAGGAGAGGAGAGGAGCTTGTGACGCTTCCGAG 604
 OY 237 TCACGTGACAGCCCTTTGCTGAGAGTGAATCTCTACGCTGAGCTGAGCTGAGAGCG 296
 DB 605 CCGTGTGACAGCCCTTCTGAGAGAGCGAGATTAACCTTAACGTATGATCTACGAGAGCG 664
 OY 297 GCGGAACAGCGCGGAGAGCTGTGCTGCTGAGAGAGCTGCTGCTGAGAGAGCTGTGAGT 356
 DB 665 GAGAAACATGTCAGAGGAGCTGTGTGTGAGAGAGCTGTGAGTGTGAGAGCGCTGTGAG 724

OY 357 GGTGATGCTGAGAGCGGCGGTGATGACAGAGCTGTGTAACGAGGCTCATGAGAGCGGCTGA 416
 DB 725 AGTCACTGTCCGCTGATGCTGTGATGATGATGAGTGTGTAATGAGTAAATGAGAGCGGACGA 784
 OY 417 CTGGGAGAGCGGCGGAGAGAGCGGCTGTGAGCTGAGCTGAGCTGAGAGAGCTGAGAGCGCT 476
 DB 785 CTGGGAGAGCTGATGAGAGAGAGCGGCTGTGAGCTGAGCTGAGAGAGCTGAGAGAGCTGAG 844
 OY 477 GCGAGCTTCTGTAACCATTAATGCTGTGAGTGTGAGAGAGCTGAGAGAGCTGAGAGCTGAG 536
 DB 845 GCGAGCTTCTGTAACCATTAATGCTGTGAGTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 904
 OY 537 CAATGACAGCTATGCTGTGAGCGCGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 596
 DB 905 CAATGACAGCTATGCTGTGAGCGCGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 964
 OY 597 CAGGCTTGGGCGCGGCTGTGAGCGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 656
 DB 965 CAGGCTTGGGCGCGGCTGTGAGCGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 1024
 OY 657 TGTGAGACTGAGAGTGTGAGAGTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 716
 DB 1025 TGTGAGACTGAGAGTGTGAGAGTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 1084
 OY 717 CTTCCTGCTGTGAGAGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 776
 DB 1085 CTTCCTGCTGTGAGAGCGGCTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 1144
 OY 777 AAGATGCTGTGAG 836
 DB 1145 AAGATGCTGTGAG 1201
 OY 837 ACAGCTTGTGAG 896
 DB 1202 ACAGCTTGTGAG 1261
 OY 897 CTTCCTGCTGTGAG 956
 DB 1262 CTTCCTGCTGTGAG 1321
 OY 957 CATGAGCGCGCTGTGAG 1016
 DB 1322 CATGAGCGCGCTGTGAG 1381
 OY 1017 TGCATGCTGTGAG 1076
 DB 1382 GGTGAGCGCGCTGTGAG 1441
 OY 1077 CCGCTGCTGTGATGATGAG 1136
 DB 1442 TCCATGCTGTGATGATGAG 1501
 OY 1137 TGTGTTGAG 1196
 DB 1502 GGTGAGCGCGCTGTGAG 1561
 OY 1197 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
 DB 1562 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1621
 OY 1257 GATGAGCGCGGAG 1313
 DB 1622 GATGAGCGCGGAG 1681
 OY 1314 ACTGAG 1373
 DB 1682 GATGAGCGCGGAG 1741
 OY 1374 GTGAG 1433
 DB 1742 GTGAG 1800
 OY 1434 CCAAGAT 1484

Mon May 19 10:50:38 2003

us-09-937-060a-19.rnpn

Page 14

Job time : 4837 secs

```
US-09-949-016-40620
: Sequence 40620, Application US/09949016
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMERISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 40620
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-40620
```

```
Query Match      38.2% Score 600.6: DB 7: Length 601:
Best Local Similarity 99.8% Pred. NO. 1.4e-122:
Matches 600: Conservative 1: Mismatches 0: Indels 0: Gaps 0:
```

```
OY 792 GACACCTGCTCCCTCCCTGTTGTGTCACAGAGGCCCGGTAGATGCACACCTTGTGCCACT 851
Db 1 GACACCTGCTCCCTCCCTGTTGTGTCACAGAGGCCCGGTAGATGCACACCTTGTGCCACT 60
OY 852 GGAGAGCCAGTGCCTCTCTACTGACAGTGTGTCGCCGACGAGACTTGTGCTAGTCTT 911
Db 61 GGAGAGCCAGTGCCTCTCTACTGACAGTGTGTCGCCGACGAGACTTGTGCTAGTCTT 120
OY 912 GGCACCTGCTGACCTGACCTGCGGAGAGATGTTGCTGACACCCATGGCCGCTGTGC 971
Db 121 GGCACCTGCTGACCTGACCTGCGGAGAGATGTTGCTGACACCCATGGCCGCTGTGC 180
OY 972 AGCTGCGTGCATGCTGTTTCTAGTGTGCGGCGGAGTGTCTGTCGTCATGCTGCTGCG 1031
Db 181 AGCTGCGTGCATGCTGTTTCTAGTGTGCGGCGGAGTGTCTGTCGTCATGCTGCTGCG 240
OY 1032 CCTCTTCTGCGCATGAGAGAGGAGGAGCATATGAGTATGATGATGATGATGATGATG 1091
Db 241 CCTCTTCTGCGCATGAGAGAGGAGGAGCATATGAGTATGATGATGATGATGATGATG 300
OY 1092 TGTGCGCGTGTGCGCTTCCGTTGAGAGCCCAAGAGTGGAAAGGTGTTGTCAGTGA 1151
Db 301 TGTGCGCGTGTGCGCTTCCGTTGAGAGCCCAAGAGTGGAAAGGTGTTGTCAGTGA 360
OY 1152 TGGGGAATTGATGCTTAGCGAGAGCCCTGACAGGCCAGTGTGACCCAACTACTTGTG 1211
Db 361 TGGGGAATTGATGCTTAGCGAGAGCCCTGACAGGCCAGTGTGACCCAACTACTTGTG 420
OY 1212 GGTGAGGGTGTGCTGAGAGCCCGCCCAAGTGAAGCCCAAGATGCCACCGCCAGA 1271
Db 421 GGTGAGGGTGTGCTGAGAGCCCGCCCAAGTGAAGCCCAAGATGCCACCGCCAGA 480
OY 1272 AGAGCCCTTATGACCCCTGAGGCGGCGGTGTCCTTAGTGTCTACTTGCAGAGACCTTCT 1331
Db 481 AGAGCCCTTATGACCCCTGAGGCGGCGGTGTCCTTAGTGTCTACTTGCAGAGACCTTCT 540
OY 1332 CCTTCCCTAGAGGCTGAGAGGCTGTCCACAGCTCTGTGAGGAGAGAGAGAGAGAGAG 1391
Db 541 CCTTCCCTAGAGGCTGAGAGGCTGTCCACAGCTCTGTGAGGAGAGAGAGAGAGAGAG 600
OY 1392 G 1392
Db 601 G 601
```

Search completed: May 17, 2003, 17:00:34

ORGANISM: Homo sapiens
US-09-937-886-9/97

Query Match 98.5%; Score 1548.8; DB 24; Length 2494;
Best Local Similarity 99.5%; Pred. No. 2.9e-298;
Matches 1544; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 GCCCCAGAGCCGCGCTGCGAGCCGCGCTGGGAGCAGCCATTAAGAGCTGAAGGCAGG 60
DB 882 GCCCCAGAGCCGCGCTGCGAGCCGCGCTGGGAGCAGCCATTAAGAGCTGAAGGCAGG 941
QY 61 AGCCGCGCGCCAGCGGCGAGCG-CCCCACAGCGCCAGGAGCCCTGGCAGCGCGGAGCCGCG 119
DB 942 AGCCGCGCGCCAGCGGCGAGCGCCCGCCACAGCGCAGGAGCCCTGGCAGCGGAGCCGCGCG 1001
QY 120 GGTGAGGATTATGATATCCACGCGGCGCGCGCGCGCGCTCTCCCGGCGCGCTGCGCGCGT 179
DB 1002 GGTGAGGATTATGATATCCACGCGGCGCGCGCGCGCGCGCTCTCCCGGCGCGCTGCGCGCGT 1061
QY 180 GCTGTGCTGCTGAACCCGCGCGCGCGCGAGGAGGAGCCCTTGCAGCTCTTCCGAGTCA 239
DB 1062 GCTGTGCTGCTGAACCCGCGCGCGCGCGAGGAGGAGCCCTTGCAGCTCTTCCGAGTCA 1121
QY 240 CGTGCAGCCCTTTTGGCTAGGCTGAATCTCTCTTACGCTGATGCTCACTGAGCGCGG 299
DB 1122 CGTGCAGCCCTTTTGGCTAGGCTGAATCTCTCTTACGCTGATGCTCACTGAGCGCGG 1181
QY 300 GAACCAAGCGCGGAGCTGCTGGGCTGCGAGAGCTGGGCGCGCGGAGCCCTTGTGTGT 359
DB 1182 GAACCAAGCGCGGAGCTGCTGGGCTGCGAGAGCTGGGCGCGCGGAGCCCTTGTGTGT 1241
QY 360 CATGTCTGAGAGCGGCGCTGATGCAAGAGTGTGTAAACGGGCTCATGAGAGCGGCTGACTG 419
DB 1242 CATGTCTGAGAGCGGCGCTGATGCAAGAGTGTGTAAACGGGCTCATGAGAGCGGCTGACTG 1301
QY 420 GAGAGCGCGCATCAAGAGCGCCCTGTGTAGCTTCCAGCAGGCTCTGGCAACGCGCTGGC 479
DB 1302 GAGAGCGCGCATCAAGAGCGCCCTGTGTAGCTTCCAGCAGGCTCTGGCAACGCGCTGGC 1361
QY 480 AGCTTCCTTGAACCATATATCTGCTGATGAGAGGTCAGCAATGAACCTCTGAGACAA 539
DB 1362 AGCTTCCTTGAACCATATATCTGCTGATGAGAGGTCAGCAATGAACCTCTGAGACAA 1421
QY 540 CTGACAGCTATTGCTGTGCGCGCGCTGCTGTACCCATGAACCTGCTGTCTGTGCACAC 599
DB 1422 CTGACAGCTATTGCTGTGCGCGCGCTGCTGTACCCATGAACCTGCTGTCTGTGCACAC 1481
QY 600 GCGTTTGGGCGCTGCGCTCTTCTGTGTGCTGACGCTGGGCTTTCATTTGCTGATGT 659
DB 1482 GCGTTTGGGCGCTGCGCTCTTCTGTGTGCTGACGCTGGGCTTTCATTTGCTGATGT 1541
QY 660 GGAAGCTAGAGTAGAGATATCGGCGCTGCGGAGAGATGGGCTGCTGCTGAGGACCT 719
DB 1542 GGAAGCTAGAGTAGAGATATCGGCGCTGCGGAGAGATGGGCTGCTGCTGAGGACCT 1601
QY 720 CCGTGTCTGTGGAGCCCTGCGCACTACCGCGCGCGAGCTGGCTTACCTCTGTAGGAG 779
DB 1602 CCGTGTCTGTGGAGCCCTGCGCACTACCGCGCGCGAGCTGGCTTACCTCTGTAGGAG 1661
QY 780 AGTGGGTTTCAAGACACTCTGCTCCCGCTTGGTGTGCTGACAGGCGCGGAGTAGTACA 839
DB 1662 AGTGGGTTTCAAGACACTCTGCTCCCGCTTGGTGTGCTGACAGGCGCGGAGTAGTACA 1721
QY 840 CTTTGTGCACTGGAGAGGAGCTGCTTCACTGCACTGAGAGTGTGCTGAGGAGCTT 899
DB 1722 CTTTGTGCACTGGAGAGGAGCTGCTTCACTGCACTGAGAGTGTGCTGAGGAGCTT 1781
QY 900 TGTGCTAGTCTGTGCACTGCTGCACTGCACTGCGAGTGTGCTGCTGAGGAGCTT 959
DB 1782 TGTGCTAGTCTGTGCACTGCTGCACTGCACTGCGAGTGTGCTGCTGAGGAGCTT 1841
QY 960 GGGCGGCTGTGAGCTGTGCTGATGCTGTCTTACGTGCGGCGGAGGTGTCTGCTGC 1019
DB 1019 GGGCGGCTGTGAGCTGTGCTGATGCTGTCTTACGTGCGGCGGAGGTGTCTGCTGC 1019

DB 1842 GGGCGGCTGTGAGCTGTGCTGATGCTGTCTTACGTGCGGCGGAGGTGTCTGCTGC 1901
QY 1020 CATGCTGCTGGCGCTTTCCTGCGCATGAGAGGAGGAGGATATGAGTATGAATGCC 1079
DB 1902 CATGCTGCTGGCGCTTTCCTGCGCATGAGAGGAGGAGGATATGAGTATGAATGCC 1961
QY 1080 CTACTTGTATATGTCGCGCTGCGCTTCCGCTTGGAGCCCAAGAGTATGGAAGTGT 1139
DB 1962 CTACTTGTATATGTCGCGCTGCGCTTCCGCTTGGAGCCCAAGAGTATGGAAGTGT 2021
QY 1140 GTTTGAGTGGAGGAGATATGATGTTAGCGAGCGCTGAGGCGAGGTGACCCAAA 1199
DB 2022 GTTTGAGTGGAGGAGATATGATGTTAGCGAGCGCTGAGGCGAGGTGACCCAAA 2081
QY 1200 CTACTTGTATGCTGAGCGGTTGCGTGTGAGAGCCCGCCCAAGCTGTGGAAGCCGAGCAT 1259
DB 2082 CTACTTGTATGCTGAGCGGTTGCGTGTGAGAGCCCGCCCAAGCTGTGGAAGCCGAGCAT 2141
QY 1260 GCCACCGCCAGAGAGCGCTTATGACCCCTGGGCGCGCGCTGCTTGTACTTGC 1319
DB 2142 GCCACCGCCAGAGAGCGCTTATGACCCCTGGGCGCGCGCTGCTTGTACTTGC 2201
QY 1320 AGGACCTTCTCTCTTCCCTTGAAGGCTGAGGCGCTGTGCACAGCTCTGTGAGGAGT 1379
DB 2202 AGGACCTTCTCTCTTCCCTTGAAGGCTGAGGCGCTGTGCACAGCTCTGTGAGGAGT 2261
QY 1380 GAGACTCTCTGCGAGAGGAGTGAAGAGTGAAGGAGCTATGCTTGGGAGGAGCAAGCCAAA 1439
DB 2262 GAGACTCTCTGCGAGAGGAGTGAAGAGTGAAGGAGCTATGCTTGGGAGGAGCAAGCCAAA 2321
QY 1440 TGAAGTCTTGGTTCAGAGAGCCAGCTGCTGGGCGCGCGCTGCTTATGAGGCGCTTGTAG 1499
DB 2322 TGAAGTCTTGGTTCAGAGAGCCAGCTGCTGGGCGCGCGCTGCTTATGAGGCGCTTGTAG 2381
QY 1500 TTTGTTCTGAGAGCCCGCCAGCCAGCAACCAATCAATTAAGTATGACATTCAGCTTGA 1559
DB 2382 TTTGTTCTGAGAGCCCGCCAGCCAGCAACCAATCAATTAAGTATGACATTCAGCTTGA 2441
QY 1560 AAAAAAAAAA 1571
DB 2442 CACCAACATGAA 2453

RESULT 8
US-09-649-163-9696
; Sequence 9696, Application US/09649163
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Galvin, Katherine A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kingsbury, Gillian A.
; APPLICANT: Welch, Nadine S.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Williamson, Mark
; APPLICANT: Richardson, Jennifer
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Villalva, Jean-Luc M.G.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Siles-Santlago, Immaculada
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; APPLICANT: Busfield, Samantha J.
; APPLICANT: Deeds, James
; APPLICANT: Lee, John
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1164-001
; CURRENT APPLICATION NUMBER: US/09/649,163
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/150,608
; PRIOR FILING DATE: 1999-08-25


```

RESULT 5
PCT-US02-24623-1
: Sequence 1, Application PC/TUS0224623
: GENERAL INFORMATION:
: APPLICANT: EXELIXIS, INC.
: TITLE OF INVENTION: SPKs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
: FILE REFERENCE: EX02-104
: CURRENT APPLICATION NUMBER: PCT/US02/24623
: PRIOR FILING DATE: 2002-08-02
: PRIOR APPLICATION NUMBER: US 60/310,362
: PRIOR FILING DATE: 2001-08-06
: PRIOR APPLICATION NUMBER: US 60/357,501
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1783
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US02-24623-1

```

```

Query Match          98.6%; Score 1550.8; DB 1: Length 1783;
Best Local Similarity 99.5%; Pred. No. 1.1e-298;
Matches 1366; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

```

QY 1 GCGCCACAGCCGCGCTGGAGCGCCGCTGGGACAGCAGCAGATAGAGAGGAGGAGG 60
DB 207 GCGCCACAGCGCGCCCTGGAGCGCCGCTGGGACAGCAGCAGATAGAGAGGAGG 266
QY 61 AGCGCGCGCGCGAGCGGAGCGG-CCGACAGCGCGCGAGGAGCGCGCGCGGAGGCGG 119
DB 267 AGCGCGCGCGCGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 326
QY 120 GCTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179
DB 327 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
QY 180 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
DB 387 GGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
QY 240 CGTGAGAGCGCGCTGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 299
DB 447 CGTGAGAGCGCGCTGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 506
QY 300 GAGCAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 359
DB 507 GAGCAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 566
QY 360 CATGCTGAGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 419
DB 567 CATGCTGAGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 626
QY 420 GAGAGCGCGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 479
DB 627 GAGAGCGCGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 686
QY 480 AGCTTCCTTGAACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 539
DB 687 AGCTTCCTTGAACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 746
QY 540 CTGACAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
DB 747 CTGACAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 806
QY 600 GCGTTCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
DB 807 GCGTTCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
QY 660 GAGCCTAGAGAGTGAAGATGAGAGTGAAGATGAGAGTGAAGATGAGAGTGAAGAT 719
DB 867 GAGCCTAGAGAGTGAAGATGAGAGTGAAGATGAGAGTGAAGATGAGAGTGAAGAT 926

```

```

QY 720 CTTGCGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 779
DB 927 CTTGCGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 986
QY 780 AGTGGCTTCCAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 839
DB 987 AGTGGCTTCCAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1046
QY 840 CTTGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 899
DB 1047 CTTGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1106
QY 900 TGTGCTAGTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 959
DB 1107 TGTGCTAGTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1166
QY 960 GGGCGCGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1019
DB 1167 GGGCGCGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1226
QY 1020 CATGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1079
DB 1227 CATGCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1286
QY 1080 CTACTGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
DB 1287 CTACTGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1346
QY 1140 GTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
DB 1347 GTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
QY 1200 CTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 1407 CTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
QY 1260 GCGCAGCGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1319
DB 1467 GCGCAGCGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1526
QY 1320 AGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
DB 1527 AGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586
QY 1380 GAGAGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1439
DB 1587 GAGAGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1646
QY 1440 TGAAGTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1499
DB 1647 TGAAGTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1706
QY 1500 TTTGTTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1559
DB 1707 TTTGTTCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1766
QY 1560 AAAAAAAAAAAAAA 1573
DB 1767 AAAAAAAAAAAAAA 1780

```

```

RESULT 6
US-09-617-081-2136
: Sequence 2136, Application US/09617081
: GENERAL INFORMATION:
: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Geering, David P.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
: TITLE OF INVENTION: HUMAN MAMMARY EPITHELIAL LIBRARY
: FILE REFERENCE: 1600.1146-001
: CURRENT APPLICATION NUMBER: US/09/617,081
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/143,929

```


OY 1500 TTTGTTCTGAGACCCCGACCCGACCAATTCGAAATTAAGTGCATTCACCCGTA 1559
 DB 2382 TTTGTTCTGAGACCCCGACCCGACCAATTCGAAATTAAGTGCATTCACCCGTA 2441
 OY 1560 AAAAAAAAAAAAAA 1573
 DB 2442 ACCAAAAAAAAAAAAA 2455

RESULT 4
 US-60-278-258-16055
 : Sequence 16055: Application US/60278258
 : GENERAL INFORMATION:
 : APPLICANT: Morris, MacDonald
 : APPLICANT: Lal, Preeti
 : APPLICANT: Diep, Binh
 : TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 : TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
 : FILE REFERENCE: GX-0010-1 P
 : CURRENT APPLICATION NUMBER: US/60/278,258
 : NUMBER OF SEQ ID NOS: 1730
 : SOFTWARE: PERL Program
 : SEQ ID NO 16055
 : LENGTH: 1991
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc.feature
 : OTHER INFORMATION: Inocyte ID No: 221818.6
 US-60-278-258-16055

Query Match 99.0%: Score 1557; DB 71; Length 1991;
 Best Local Similarity 99.98; Pred. No. 6.5e-300;
 Matches 1568; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GCCCCACAGCCGCGCCCTGAGCGCCGCTGGGAGAGCCGATGAGAGCTGAAGGACAG 60
 DB 423 GCCCCACAGCCGCGCCCTGAGCGCCGCTGGGAGAGCCGATGAGAGCTGAAGGACAG 482
 OY 61 AGCCGCGCGCCAGGAGAGG--CCCCACAGCCGCGAGAGCCGCTGGGAGAGGCGG 119
 DB 483 AGCCGCGCGCCAGGAGAGGCGCCGCGAGAGCCGCGAGAGCCGCTGGGAGAGGCGG 542
 OY 120 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 179
 DB 543 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 602
 OY 180 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 239
 DB 603 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 662
 OY 240 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 299
 DB 663 GGTGAGGTTATGATGATCCAGCGGCGGCGCCGCGGCGTGTCCGCGCGCCGCTGG 722
 OY 300 GAACCAAGCGCGGAGCTGTGCGGTGAGAGAGCTGGGCGCGCTGGAGCTGTGGT 359
 DB 723 GAACCAAGCGCGGAGCTGTGCGGTGAGAGAGCTGGGCGCGCTGGAGCTGTGGT 782
 OY 360 CATGTCTGAGAGCGGAGCTGTGAGAGAGCTGGAGAGCTGGAGAGCTGGAGAGCTG 419
 DB 783 CATGTCTGAGAGCGGAGCTGTGAGAGAGCTGGAGAGCTGGAGAGCTGGAGAGCTG 842
 OY 420 GAGAGCGCGCATTCAGAGAGCCGCTGTAGCTTCAGAGAGCTGTGGCAAGCGCGTGG 479
 DB 843 GAGAGCGCGCATTCAGAGAGCCGCTGTAGCTTCAGAGAGCTGTGGCAAGCGCGTGG 902
 OY 480 AGCTCTCTGAGAGCATTCAGAGAGCTGTGAGAGAGCTGTGGCAAGCGCGTGGCA 539
 DB 903 AGCTCTCTGAGAGCATTCAGAGAGCTGTGAGAGAGCTGTGGCAAGCGCGTGGCA 962

OY 540 CTGACAGCTATTGCTGTGCGCGCGGCGTGTGTCACCCATGAACTGTGTCTGCACAC 599
 DB 963 CTGACAGCTATTGCTGTGCGCGCGGCGTGTGTCACCCATGAACTGTGTCTGCACAC 1022
 OY 600 GGGTGGGGGCGGCGGCGCTTCTGTGCTCAGGCTGGGCTTCAATGCTGATCT 659
 DB 1023 GGGTGGGGGCGGCGGCGCTTCTGTGCTCAGGCTGGGCTTCAATGCTGATCT 1082
 OY 660 GGAAGCTAGAGAGTGAAGATGCGGCTGTGGGAGATGCGCTTCACTGTGGCAGCTT 719
 DB 1083 GGAAGCTAGAGAGTGAAGATGCGGCTGTGGGAGATGCGCTTCACTGTGGCAGCTT 1142
 OY 720 CCGGCGTGTGAGCGCCCTGCGCACTTACCGCGCGCAGCTGCGCTTCACTGTGAGAG 779
 DB 1143 CCGGCGTGTGAGCGCCCTGCGCACTTACCGCGCGCAGCTGCGCTTCACTGTGAGAG 1202
 OY 780 AGTGGCTTCAAGACACCTGCTCCCGCTTGTGTGTCAGAGAGGCGCGGATGATGACA 839
 DB 1203 AGTGGCTTCAAGACACCTGCTCCCGCTTGTGTGTCAGAGAGGCGCGGATGATGACA 1262
 OY 840 CCTTGTGACAGTGGAGAGAGCGCCCTCTCACTGAGACAGTGGCGCGAGAGACTT 899
 DB 1263 CCTTGTGACAGTGGAGAGAGCGCCCTCTCACTGAGACAGTGGCGCGAGAGACTT 1322
 OY 900 TGTGCTAGTCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 959
 DB 1323 TGTGCTAGTCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1382
 OY 960 GGGCGGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1019
 DB 1383 GGGCGGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1442
 OY 1020 CATGCTGTGCGGCTCTTCTGCTGCGCATGAGAGAGGCGAGCATATGATGATGCCCC 1079
 DB 1443 CATGCTGTGCGGCTCTTCTGCTGCGCATGAGAGAGGCGAGCATATGATGATGCCCC 1502
 OY 1080 CTACTGTGATATGTCGCGCGTGTGTCGCGCTTCCGTTGAGAGCCCAAGATGAGGAA 1139
 DB 1503 CTACTGTGATATGTCGCGCGTGTGTCGCGCTTCCGTTGAGAGCCCAAGATGAGGAA 1562
 OY 1140 GTTGTGAGTGAATGAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1199
 DB 1563 GTTGTGAGTGAATGAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 1622
 OY 1200 CTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 DB 1623 CTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1682
 OY 1260 GCGACGCGCAGAGAGCGCTTATGACCCCTGCGGCGCGCTGTGCTTATGCTCTAC 1319
 DB 1683 GCGACGCGCAGAGAGCGCTTATGACCCCTGCGGCGCGCTGTGCTTATGCTCTAC 1742
 OY 1320 AGGACGCTTCTCTCTTCTGATGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAG 1379
 DB 1743 AGGACGCTTCTCTCTTCTGATGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAG 1802
 OY 1380 GAGACTCTCTGAGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAG 1439
 DB 1803 GAGACTCTCTGAGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAG 1862
 OY 1440 TGAAGTCTGTGAGTGAAGAGCCAGCTGTGAGGCGAGCTGTGAGTGAAGGCTGTGAG 1499
 DB 1863 TGAAGTCTGTGAGTGAAGAGCCAGCTGTGAGGCGAGCTGTGAGTGAAGGCTGTGAG 1922
 OY 1500 TTTGTTCTGAGAGCCCGCCAGCAAGCAATTCGAAATTAAGTGAATTCGAGCCTGA 1559
 DB 1923 TTTGTTCTGAGAGCCCGCCAGCAAGCAATTCGAAATTAAGTGAATTCGAGCCTGA 1982
 OY 1560 AAAAAAAAAA 1568
 DB 1983 AAAAAAAAAA 1991

```

|||||
Db 1381 AGACCTCTGTGAGAGAGGAGGAGAGAGCTGAGAGCTATGCTTTGGGGGAGAGCCAGAAAT 1440
|||
QY 1441 GAAGCTCTGGGTGAGAGAGCCAGCTGGGGGCGGCTGAGTGAAGAGCCCTTACAT 1500
|||||
Db 1441 GAAGTCTGGGTGAGAGAGCCAGCTGGGGGCGGCTGAGTGAAGAGCCCTTACAT 1500
|||||
QY 1501 TTGTTCTGAGAGCCAGCCAGCAACCAATCAATAAGTACATTTCCAGCTGAA 1560
|||||
Db 1501 TTGTTCTGAGAGCCAGCCAGCAACCAATCAATAAGTACATTTCCAGCTGAA 1560
|||||
QY 1561 AAAAAAAAAA 1573
|||||
Db 1561 AAAAAAAAAA 1573

RESULT 3
US-09-526-996-1216
; Sequence 1216, Application US/09526996
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM
; TITLE OF INVENTION: HUMAN PROSTATE AND BREAST LIBRARIES
; FILE REFERENCE: 1600.1088-001
; CURRENT APPLICATION NUMBER: US/09/526,996
; EARLIER FILING DATE: 2000-03-15
; EARLIER APPLICATION NUMBER: 60/125,469
; EARLIER FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 60/125,560
; NUMBER OF SEQ ID NOS: 1296
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1216
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2494)
; OTHER INFORMATION: n = A,T,C or G
US-09-526-996-1216

Query Match 99.0%; Score 1557.2; DB 19; Length 2494;
Best Local Similarity 99.7%; Pred. No. 6,1e-300;
Matches 1570; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 420 GGAGACCGCATCCAGAGCCCTGTGTACCTCCAGAGAGGCTTGGCAAGCGGTGC 479
|||||
Db 1302 GGAGACCGCATCCAGAGCCCTGTGTACCTCCAGAGAGGCTTGGCAAGCGGTGC 1361
|||||
QY 480 AGCTTCCTTGAAACCAATTATGCTGGCTATGAGAGAGTACCAATGAAGACCTCTGACCA 539
|||||
Db 1362 AGCTTCCTTGAAACCAATTATGCTGGCTATGAGAGAGTACCAATGAAGACCTCTGACCA 1421
|||||
QY 540 CTGACAGCTATTTCTGTGGCGCGGCTGCTGTACACCATTAACCTGCTCTGTGACAC 599
|||||
Db 1422 CTGACAGCTATTTCTGTGGCGCGGCTGCTGTACACCATTAACCTGCTCTGTGACAC 1481
|||||
QY 600 GGCTTCGGGCTCGGCTCTTCTGTGCTGACCTGGAGCTTGGAGCTTATGCTATGT 659
|||||
Db 1482 GGCTTCGGGCTCGGCTCTTCTGTGCTGACCTGGAGCTTGGAGCTTATGCTATGT 1541
|||||
QY 660 GGACCTAGAGAGTGAAGATGAGGCTTGGGGGAGATGAGGCTTACCTGTGGCACTT 719
|||||
Db 1542 GGACCTAGAGAGTGAAGATGAGGCTTGGGGGAGATGAGGCTTACCTGTGGCACTT 1601
|||||
QY 720 CTGCGCTGTGGAGCCCTGCGCACCTACCGCGGCGGCTGCTTACCTCTGTAGAG 779
|||||
Db 1602 CTGCGCTGTGGAGCCCTGCGCACCTACCGCGGCGGCTGCTTACCTCTGTAGAG 1661
|||||
QY 780 AGTGGGTCCAGAGACACCTGCTCCCGCGGCTTGGGTGTCACAGAGCGGCTAGATGACA 839
|||||
Db 1662 AGTGGGTCCAGAGACACCTGCTCCCGCGGCTTGGGTGTCACAGAGCGGCTAGATGACA 1721
|||||
QY 840 CCTTGTCCACTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 899
|||||
Db 1722 CCTTGTCCACTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
|||||
QY 900 TGTGCTAGTCTGTGGCACTGCTGCACTGCGACCTGGGAGTGAATGTTTCTGTGACCCAT 959
|||||
Db 1782 TGTGCTAGTCTGTGGCACTGCTGCACTGCGACCTGGGAGTGAATGTTTCTGTGACCCAT 1841
|||||
QY 960 GGGCGGCTGTGACAGTGGGCTGATGATGCTGTTCTACGTGCGGCGGAGATGCTGTGTC 1019
|||||
Db 1842 GGGCGGCTGTGACAGTGGGCTGATGATGCTGTTCTACGTGCGGCGGAGATGCTGTGTC 1901
|||||
QY 1020 CATGCTGCTGCGCTTCTTCTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079
|||||
Db 1902 CATGCTGCTGCGCTTCTTCTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1961
|||||
QY 1080 CTACTTGTATATGTGCGCGCTGCTGCGCTTCTGCGCATGAGAGGAGGAGGAGGAGGAGG 1139
|||||
Db 1962 CTACTTGTATATGTGCGCGCTGCTGCGCTTCTGCGCATGAGAGGAGGAGGAGGAGGAGG 2021
|||||
QY 1140 GTTGTGAGTGGATGGGAGATTTGATGCTTACGAGGCTGTCAGGCGGAGGAGGAGGAGG 1199
|||||
Db 2022 GTTGTGAGTGGATGGGAGATTTGATGCTTACGAGGCTGTCAGGCGGAGGAGGAGGAGG 2081
|||||
QY 1200 CTACTTGTGATGCTGAGCGTGTGCGTGGAGAGGCGGCGGAGGAGGAGGAGGAGGAGG 1259
|||||
Db 2082 CTACTTGTGATGCTGAGCGTGTGCGTGGAGAGGCGGCGGAGGAGGAGGAGGAGGAGG 2141
|||||
QY 1260 GCCACCGCCAGAGAGGCGCTTATGACCCCTGGGCGGCTGCTGCTTATGCTTACCTTGC 1319
|||||
Db 2142 GCCACCGCCAGAGAGGCGCTTATGACCCCTGGGCGGCTGCTGCTTATGCTTACCTTGC 2201
|||||
QY 1320 AGGACCTTCTCTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1379
|||||
Db 2202 AGGACCTTCTCTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2261
|||||
QY 1380 GAGACTCTCTGAGAGAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1439
|||||
Db 2262 GAGACTCTCTGAGAGAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 2321
|||||
QY 1440 TGAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1499
|||||
Db 2322 TGAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2381

```

Oy	1261	GCACCGCAGAGACCCCTTTGACACTGAGCGCGGCGTGGTACCTTACTTGTCA	1320
Db	1261	CCACCGCAGAGACCCCTTTGACCCCTGGCGCGGCGTGGTACCTTACTTGTCA	1320
Oy	1321	GCACCTTCTCTCTCCCTTAGCGCTGAGGCGCTGTGCACAGCTCTGTGGGGGTGGAG	1380
Db	1321	GCACCTTCTCTCTCTCCCTTAGCGCTGAGGCGCTGTGCACAGCTCTGTGGGGGTGGAG	1380
Oy	1381	AGACCTCTGTGAGAAAGGTGAGAGGTGAGCGTATGCTTTGGGGGACAGGCGCGAAT	1440
Db	1381	AGACCTCTGTGAGAAAGGTGAGAGGTGAGCGTATGCTTTGGGGGACAGGCGCGAAT	1440
Oy	1441	GAAGTCTCTGGGTCAAGAGACCACAGTGGCTGGCGCCAGCTGCCCTATGTAAAGGCTTTAGT	1500
Db	1441	GAAGTCTCTGGGTCAAGAGACCACAGTGGCTGGCGCCAGCTGCCCTATGTAAAGGCTTTAGT	1500
Oy	1501	TTGTGTTCTGAGACCCCCACCCACGACGAAACCAATTCACAAATTAAGTGAATTCACAGCTGAA	1560
Db	1501	TTGTGTTCTGAGACCCCCACCCACGACGAAACCAATTCACAAATTAAGTGAATTCACAGCTGAA	1560
Oy	1561	AAAAAAAAAAAAA 1573	
Db	1561	AAAAAAAAAAAAA 1573	

RESULT 2
US-60-125-593-12
: Sequence 12, Application US/60125593

```

? APPLICANT: Bandman, Olga
? APPLICANT: Tang, Y. Tom
? APPLICANT: Yue, Henry
? APPLICANT: Baughn, Mariah R.
? APPLICANT: Azimzal, Yalda
? TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
? FILE REFERENCE: PF-0683 P
? CURRENT APPLICATION NUMBER: US/60/125,593
? CURRENT FILING DATE: 1999-03-18
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PERL Program
? SEQ ID NO 12
? LENGTH: 1573
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: 2415617
US-60-125-593-12

```

Query Match	100.0%;	Score 1573;	DB 56;	Length 1573;
Best Local Similarity	100.0%;	Pred. No. 4.1e-303;		
Matches 1573;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	GGCCCAAGCGGAGCCCTGGGAGCGCCCGCTGGGACACCGCTAAGAGCTGAAGCGAGG	60
Db	1	GGCCCAAGCGGAGCCCTGGGAGCGCCCGCTGGGACACCGCTAAGAGCTGAAGCGAGG	60
OY	61	AGCCCGCCCAAGGGGAGCGGCCCAAGCGCAGAGGACCCCTGGCAAGGGAGGAGCGGGG	120
Db	61	AGCCCGCCCAAGGGGAGCGGCCCAAGCGCAGAGGACCCCTGGCAAGGGAGGAGCGGGG	120
OY	121	GTGCAGGTTATGATCATCAGAGGGGCGCGCCCGCGGGCGTGTCCGCGCGCTTCGCGCGTG	180
Db	121	GTGCAGGTTATGATCATCAGAGGGGCGCGCCCGCGGGCGTGTCCGCGCGCTTCGCGCGTG	180
OY	181	CTGTGTCTCTCTGAACCCGCGCGCGGGCGCAAGGCGCTTTCAGAGCTTTCGCGAGGTAC	240
Db	181	CTGTGTCTCTCTGAACCCGCGCGCGGGCGCAAGGCGCTTTCAGAGCTTTCGCGAGGTAC	240
OY	241	GTGCAGCCCCCTTTGGCTGAGGCTGTAAATCTCTTACCTGATGCTCACTAGAGCGGGG	300
Db	241	GTGCAGCCCCCTTTGGCTGAGGCTGTAAATCTCTTACCTGATGCTCACTAGAGCGGGG	300

QY	301	AACCAAGCCCGGAGACTCTGTCGCGTGGAAAGACTGGGCGCGTGGAGCGCTGAGTGGTCTC	360
Dp	301	AACCAAGCCCGGAGACTCTGTCGCGTGGAAAGACTGGGCGCGTGGAGCGCTGAGTGGTCTC	360
QY	361	ATGTCGAGAGAGGGGCTGATGACAGAGGTGGTGAACGGGGCTCATGGAGCGGGCTGGCTGG	420
Dp	361	ATGTCGAGAGAGGGGCTGATGACAGAGGTGGTGAACGGGGCTCATGGAGCGGGCTGGCTGG	420
QY	421	GAAACGGGCATCCAGAAAGCCCTGTGTAGCTTCCACGAGGCTCTGGCAACGGCGGTGTGTA	480
Dp	421	GAAACGGGCATCCAGAAAGCCCTGTGTAGCTTCCACGAGGCTCTGGCAACGGCGGTGTGTA	480
QY	481	GCCTTCCTTGAACCATTTATGCTGGCTATGAGCAGGTGACCAATGAAAGACTCTGACCAAC	540
Dp	481	GCCTTCCTTGAACCATTTATGCTGGCTATGAGCAGGTGACCAATGAAAGACTCTGACCAAC	540
QY	541	TGCACGCTATTGCTGTGGCGCGGCGGCTGTGCTACCCATGTAAACCTGTGTCTGTGCACAG	600
Dp	541	TGCACGCTATTGCTGTGGCGCGGCGGCTGTGCTACCCATGTAAACCTGTGTCTGTGCACAG	600
QY	601	GCCTTCGAGGCTCGGCGCTCTTCTCTGTGTGCTCAGCTGTGGCGCTTGACATGTGTGATGTG	660
Dp	601	GCCTTCGAGGCTCGGCGCTCTTCTCTGTGTGCTCAGCTGTGGCGCTTGACATGTGTGATGTG	660
QY	661	GACCTAGAGATGAGAAATATACGGCGTGTGGGGAGATAGCTTCACTCTGAGGCACTTTC	720
Dp	661	GACCTAGAGATGAGAAATATACGGCGTGTGGGGAGATAGCTTCACTCTGAGGCACTTTC	720
QY	721	CTGCGCTGTGGAGCGCCCTGCGACCTTACCGGGGGGACATAGGGCTACCTCCCTGTAGGAGAA	780
Dp	721	CTGCGCTGTGGAGCGCCCTGCGACCTTACCGGGGGGACATAGGGCTACCTCCCTGTAGGAGAA	780
QY	781	GTTGGGTTCCAGAAACCGGCGTCCCGGTGTGTGTGCAGAGAGGCGCGGTAGATGTAGACAC	840
Dp	781	GTTGGGTTCCAGAAACCGGCGTCCCGGTGTGTGTGCAGAGAGGCGCGGTAGATGTAGACAC	840
QY	841	CTTGTGTCCACTGTGAGAGCCAGTGTGCTCTCACTGTGACAGTGTGTGCCGAGAGAGACTTT	900
Dp	841	CTTGTGTCCACTGTGAGAGCCAGTGTGCTCTCACTGTGACAGTGTGTGCCGAGAGAGACTTT	900
QY	901	GTTCTAGTCTGTGGCACTCTGTCTACATGTGACACTGTGGGAGATGATTTGTGTGACCCATG	960
Dp	901	GTTCTAGTCTGTGGCACTCTGTCTACATGTGACACTGTGGGAGATGATTTGTGTGACCCATG	960
QY	961	GGCCGCTGTGACGTCGGGCGTCATGTCTGTCTACGTGCGGGGGGAGATGTCTCTGCC	1020
Dp	961	GGCCGCTGTGACGTCGGGCGTCATGTCTGTCTACGTGCGGGGGGAGATGTCTCTGCC	1020
QY	1021	ATGTCCTGTCCGCTCTTCTGTGGCATTGTAGAGAAAGGCGAGCATATGTGAGTATGTAATCCCC	1080
Dp	1021	ATGTCCTGTCCGCTCTTCTGTGGCATTGTAGAGAAAGGCGAGCATATGTGAGTATGTAATCCCC	1080
QY	1081	TACTTGTATATGTATGTCGCGTGTGTCGCGCTTCCGTTGTGAGACCCAAAGATGGGAAAGTGTG	1140
Dp	1081	TACTTGTATATGTATGTCGCGTGTGTCGCGCTTCCGTTGTGAGACCCAAAGATGGGAAAGTGTG	1140
QY	1141	TTTTCGAGTGTGGGGAATATGATGTATGACGAGGCGGTGCAGAGGCGCAGGTGCACCCCAAC	1200
Dp	1141	TTTTCGAGTGTGGGGAATATGATGTATGACGAGGCGGTGCAGAGGCGCAGGTGCACCCCAAC	1200
QY	1201	TACTTGTGATGTGTACAGGT	1260
Dp	1201	TACTTGTGATGTGTACAGGT	1260
QY	1261	CCACGCGCAGAAAGACCCCTTATGACCCCTGTGGCGCGCTGTGCTTATGTCTTACTTTGCA	1320
Dp	1261	CCACGCGCAGAAAGACCCCTTATGACCCCTGTGGCGCGCTGTGCTTATGTCTTACTTTGCA	1320
QY	1321	GGAACCTTCTCTCTTCCCTATGAGGCGTGTACAGGGCTGTGCACAGCTCTGTGTGGGGTGTGAGC	1380
Dp	1321	GGAACCTTCTCTCTCTTCCCTATGAGGCGTGTACAGGGCTGTGCACAGCTCTGTGTGGGGTGTGAGC	1380
QY	1381	AGACCTCTCTGTGAGAGGGTGTAGAAAGTGTGAGAGCTATGCTTTTGGGGGAGAGGCGCCAGAT	1440
Dp	1381	AGACCTCTCTGTGAGAGGGTGTAGAAAGTGTGAGAGCTATGCTTTTGGGGGAGAGGCGCCAGAT	1440

```
22 1539.4 97.9 1772 24 US-09-629-469A-15393 Sequence 15393, A
23 1539 97.8 1719 40 US-10-111-729-1 Sequence 1, Appl1
24 1524.6 96.9 1797 17 US-09-338-467-876 Sequence 876, App
25 1524.6 96.9 1797 31 US-09-808-384-876 Sequence 876, App
26 1524.6 96.9 1797 31 US-09-813-206-806 Sequence 806, App
27 1515.4 96.3 1859 20 US-09-532-805-118 Sequence 118, App
28 1511.6 96.1 1600 30 US-09-784-810A-1 Sequence 1, Appl1
29 1479.2 94.0 1533 1 PCT-US01-05614-90 Sequence 90, Appl1
30 1479.2 94.0 1533 1 PCT-US98-11422A-90 Sequence 90, Appl1
31 1479.2 94.0 1533 16 US-09-205-258-90 Sequence 90, Appl1
32 1479.2 94.0 1533 35 US-09-933-767-90 Sequence 90, Appl1
33 1479.2 94.0 1533 38 US-10-004-860-90 Sequence 90, Appl1
34 1479.2 94.0 1533 38 US-10-023-282-90 Sequence 90, Appl1
35 1430.2 90.9 1562 1 PCT-US01-16450A-1180 Sequence 1180, Ap
36 1430.2 90.9 1562 1 PCT-US01-16450A-1180 Sequence 1180, Ap
37 1428.4 90.8 1338 18 US-09-488-725A-1539 Sequence 1539, Ap
38 1428.4 90.8 1438 25 US-09-653-450A-129 Sequence 129, App
39 1258.4 80.0 1272 1 PCT-US02-24623-2 Sequence 2, Appl1
40 1168.6 74.3 1205 36 US-09-959-897-1 Sequence 1, Appl1
41 1165 74.1 1173 80 US-60-364-517-1 Sequence 1, Appl1
42 1163.2 73.9 1200 18 US-09-471-275-5302 Sequence 5302, Ap
43 1163.2 73.9 1200 18 US-09-488-725A-5111 Sequence 5111, Ap
44 1155 73.4 1155 36 US-09-970-516-1 Sequence 1, Appl1
45 1155 73.4 1155 40 US-10-111-729-2 Sequence 2, Appl1
```

ALIGNMENTS

RESULT 1

US-09-937-060a-19

Sequence 19, Application US/09937060A

GENERAL INFORMATION:

APPLICANT: INCYTE PHARMACEUTICALS, INC.

APPLICANT: BANDMAN, Olga

APPLICANT: TANG, Y. Tom

APPLICANT: YUE, Henry

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: AZIMZAI, Yalda

APPLICANT: LU, Dying Anna M.

APPLICANT: AU-YOUNG, Janice

TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION

FILE REFERENCE: PF-0683 PCT

CURRENT APPLICATION NUMBER: US/09/937,060A

PRIOR FILING DATE: 2002-04-15

PRIOR APPLICATION NUMBER: 60/125,593; 60/135,049; 60/143,188

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PERL Program

SEQ ID NO 19

LENGTH: 1573

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 2415617CB1

US-09-937-060a-19

```
Query Match 100.0%; Score 1573; DB 35; Length 1573;
Best Local Similarity 100.0%; Pred. No. 4,1e+303;
Matches 1573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCCCACAGCGCGCTGCGACGCCGCTGGGACGACCGATPAGAGCTGAAGCAGG 60
DB 1 GCCCACAGCGCGCTGCGACGCCGCTGGGACGACCGATPAGAGCTGAAGCAGG 60
QY AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GTGAGGTTATGATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 121 GTGAGGTTATGATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
```

```
DB 121 GTGAGGTTATGATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 CTGTGTCTGTGTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 181 CTGTGTCTGTGTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 GTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 GTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 AACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 AACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 ATGTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 ATGTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 GAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 GAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 GCTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 GCTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 TGCAGCGCTATGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 TGCAGCGCTATGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 GCTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 601 GCTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 661 GACCTGAGAGAGTGAAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 GACCTGAGAGAGTGAAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 CTGCGTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 CTGCGTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 GTGAGGTTCCAGACACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 GTGAGGTTCCAGACACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 CTGTGTGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 CTGTGTGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GTGTGTGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 901 GTGTGTGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 961 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1021 ATGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 ATGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 TACTTGTATATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1081 TACTTGTATATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 TTTGAGATGATGAGGAAATGATGTTAGGAGGCGCGCGCGCGCGCGCGCG 1200
DB 1141 TTTGAGATGATGAGGAAATGATGTTAGGAGGCGCGCGCGCGCGCGCGCG 1200
QY 1201 TACTTGTATGATGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
DB 1201 TACTTGTATGATGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
```



```

OY 190 MRETLGTLRLAALRTYRGLAYLVPG-RVGSKTPA----- 224
      ||||| ||| ||||| |||
Db 297 ARFTLGTGLATLHTYRGLSYLPATVEPASPTPAHSI.PRAKSELTLTPDPAPMAHSP 356
OY 225 ----- 224
Db 357 LHRVSDDLPLPLPAPALASPSHEPLPLLSLNGGPELAGDMGAGADAPLSPDLLSSPP 416
OY 225 -----SPVVVOO-----GPDAAHLVPLEBPVSHMTVVP 253
      ||||| ||| ||||| |||
Db 417 GSPKALHSPVSGAVIPSSSGLPLPTPDARVGASTCCPPDHLPLPLGTPPPDM-VTL 475
OY 254 DEDFVLVLLHSHLSEMFAPMGRACAGVMHLFFYRAGVSRAMLRLFLAMEKGRHNE 313
      ||||| ||| ||||| |||
Db 476 EGDFTVLMALISPSHLCADLVAAHAPARDGDLVHLCWVRSGISRAALLRLFLAMERSHFS 535
OY 314 YECPLYVYVVAARFLEPKDKGVAVDGLMSEAVOGVHPNYPMWVSGCVERPPSK 373
      ||||| ||| ||||| |||
Db 536 LGGPOLGYAARAFRLPLRGVLTVDGEVGYPLQOMHPGIGTLTLCG---PPGC- 590
OY 374 PQOMPEEP 383
      |||
Db 591 -----PGRFP 595

RESULT 12
PCT-US03-02588-78
: Sequence 78, Application PC/TUS0302588
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc
: APPLICANT: Hunter, John Joseph
: APPLICANT: Macbeth, Kyle J.
: APPLICANT: Tsai, Fong-Ying
: APPLICANT: Lesoon, Andrea
: APPLICANT: Lightcap, Eric S.
: APPLICANT: Williamson, Mark
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
: TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
: TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
: TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
: TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
: TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
: TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
: TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
: FILE REFERENCE: MP102-020
: CURRENT APPLICATION NUMBER: PCT/US03/02588
: CURRENT FILING DATE: 2003-01-30
: PRIOR APPLICATION NUMBER: US 60/353,600
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 60/364,517
: PRIOR FILING DATE: 2002-03-15
: PRIOR APPLICATION NUMBER: US 60/371,075
: PRIOR FILING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: US 60/371,507
: PRIOR FILING DATE: 2002-04-10
: PRIOR APPLICATION NUMBER: US 60/372,984
: PRIOR FILING DATE: 2002-04-16
: PRIOR APPLICATION NUMBER: US 60/374,194
: PRIOR FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: US 60/382,995
: PRIOR FILING DATE: 2002-05-24
: PRIOR APPLICATION NUMBER: US 60/385,023
: PRIOR FILING DATE: 2002-05-31
: PRIOR APPLICATION NUMBER: US 60/388,853
: PRIOR FILING DATE: 2002-06-14
: PRIOR APPLICATION NUMBER: US 60/389,395
: PRIOR FILING DATE: 2002-06-17
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 122
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 78
: LENGTH: 618

```

```

: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US03-02588-78
Query Match 44.0%; Score 886.5; DB 1: Length 618;
Best Local Similarity 39.8%; Pred. No. 1,6e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

OY 10 VLPRPCRVILNIPRCGKKALQLOFRSHYQPLLAEEISFTLMLTPRRHARELYSEEL 69
      ||||| ||| ||||| |||
Db 140 LDRPRLDLVNPFGGRLAMOWCNHVLPMISEGLSFNLIOTERQNHARELVGCLSL 199
OY 70 GRDVLVNSGDLGMEVYVNGLMERPDMETATQKPLCSIPAGSNAALASLNHYAGYEV 129
      ||| ||||| ||||| |||
Db 200 SEMDLITVYSGDLHEVLNGLIDRPDMEEAVKMPYGIILPCGSNAALAGAVNQGGEPA 259
OY 130 TNEDLITNCTLLCRLLSPMLLSLHTYASGLRFSVLSLACGFADVDLESKTRRICE 189
      ||| ||||| ||||| |||
Db 260 LGLDILLNCSLLRCGGGRLDLSVTLASGSRCSFSLVAMGFVDVDIQSERFALS 319
OY 190 MRETLGTLRLAALRTYRGLAYLVPG-RVGSKTPA----- 224
      ||||| ||| ||||| |||
Db 320 ARFTLGTGLATLHTYRGLSYLPATVEPASPTPAHSI.PRAKSELTLTPDPAPMAHSP 379
OY 225 ----- 224
Db 380 LHRVSDDLPLPLPAPALASPSHEPLPLLSLNGGPELAGDMGAGADAPLSPDLLSSPP 439
OY 225 -----SPVVVOO-----GPDAAHLVPLEBPVSHMTVVP 253
      ||||| ||| ||||| |||
Db 440 GSPKALHSPVSGAVIPSSSGLPLPTPDARVGASTCCPPDHLPLPLGTPPPDM-VTL 498
OY 254 DEDFVLVLLHSHLSEMFAPMGRACAGVMHLFFYRAGVSRAMLRLFLAMEKGRHNE 313
      ||||| ||| ||||| |||
Db 499 EGDFTVLMALISPSHLCADLVAAHAPARDGDLVHLCWVRSGISRAALLRLFLAMERSHFS 558
OY 314 YECPLYVYVVAARFLEPKDKGVAVDGLMSEAVOGVHPNYPMWVSGCVERPPSK 373
      ||||| ||| ||||| |||
Db 559 LGGPOLGYAARAFRLPLRGVLTVDGEVGYPLQOMHPGIGTLTLCG---PPGC- 613
OY 374 PQOMPEEP 383
      |||
Db 614 -----PGRFP 618

RESULT 13
US-10-354-358-78
: Sequence 78, Application US/10354358
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc
: APPLICANT: Hunter, John Joseph
: APPLICANT: Macbeth, Kyle J.
: APPLICANT: Tsai, Fong-Ying
: APPLICANT: Lesoon, Andrea
: APPLICANT: Lightcap, Eric S.
: APPLICANT: Williamson, Mark
: APPLICANT: Rudolph-Owen, Laura A.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
: TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
: TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
: TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
: TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
: TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
: TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
: TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULE
: FILE REFERENCE: MP102-020PLRNMNM
: CURRENT APPLICATION NUMBER: US/10/354,358
: CURRENT FILING DATE: 2003-01-30
: PRIOR APPLICATION NUMBER: US 60/353,600
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 60/364,517
: PRIOR FILING DATE: 2002-03-15
: PRIOR APPLICATION NUMBER: US 60/371,075

```

Mon May.19 10:50:39 2003

us-09-937-060a-5.1a.pn

Page 5

CURRENT APPLICATION NUMBER: US/60/452.680
CURRENT FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 116213
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12558
LENGTH: 595
TYPE: PRT
ORGANISM: Homo sapiens
US-60-452-680-12558

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1.5e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

OY 10 VLRPCRVLLVLPNGCKKALQLEFRSHVOPLLAEAFISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLRLLLVNPFGGGLAMQCKNHNVLPMISEAGLSFNLQTERNNHARELVQGLSL 176
OY 70 GRMDALVYVSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 129
DB 177 SEMDGIYVYSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 236
OY 130 TNEDLTNGTLLCRRLSPMNLISLHTASGLRFLSVSLAMGFADVDLESEKRYRLGE 189
DB 237 LGIDLNLNCSLLCCGGCHPLDLISVTLASGSCFSLVAMGFVSDVDIOSEKFRALGS 296
OY 190 MRFTLGTLRLAALRTYGRGLAYLPVG-RVGSKTPA----- 224
DB 297 ARPTLGTVGLATLHTYGRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 356
OY 225 ----- 224,
DB 357 LHRVSDLPPLPOPALASPGSPERPLPLSLNGGPELAGDMGAGADAPLSDDLSSPP 416
OY 225 -----SPVYVQO-----GPVDAHLVPLEEVPSPHMTYVP 253
DB 417 GSPKALHSFVSEGAAPVLPSSGLPLPTPDARVCASTGCPDHLPLPGTLPDPDW-VTL 475
OY 254 DEDFVLYALHSHIGSEMFAAPMRCACAGVHMLFYVRAGVSRAMLRLFLAMEGRHME 313
DB 476 ECGFVLMALISPSHIGADLVAAPIARDDQVHLCVNRGSISSALRLFLAMEGRSHFS 535
OY 314 YECPLYVYVVAARFLERKDGKGFANVDSGLMSEAVNGQVPTFMVNSGVCEPPPSMK 373
DB 536 LCGFOLGYAAARAFRLERPLTRGVLYVDGEVYGPLOAMHPIGTLTGTG---PCC- 590
OY 374 PQMPPEEP 383
DB 591 -----PGRPEP 595

RESULT 10

US-60-453-135-7837
Sequence 7837, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARCILL, Michele
APPLICANT: JAKOBYA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO01456
CURRENT APPLICATION NUMBER: US/60/453.135
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7837
LENGTH: 595
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-135-7837

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1.5e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

OY 10 VLRPCRVLLVLPNGCKKALQLEFRSHVOPLLAEAFISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLRLLLVNPFGGGLAMQCKNHNVLPMISEAGLSFNLQTERNNHARELVQGLSL 176
OY 70 GRMDALVYVSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 129
DB 177 SEMDGIYVYSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 236
OY 130 TNEDLTNGTLLCRRLSPMNLISLHTASGLRFLSVSLAMGFADVDLESEKRYRLGE 189
DB 237 LGIDLNLNCSLLCCGGCHPLDLISVTLASGSCFSLVAMGFVSDVDIOSEKFRALGS 296
OY 190 MRFTLGTLRLAALRTYGRGLAYLPVG-RVGSKTPA----- 224
DB 297 ARPTLGTVGLATLHTYGRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 356
OY 225 ----- 224
DB 357 LHRVSDLPPLPOPALASPGSPERPLPLSLNGGPELAGDMGAGADAPLSDDLSSPP 416
OY 225 -----SPVYVQO-----GPVDAHLVPLEEVPSPHMTYVP 253
DB 417 GSPKALHSFVSEGAAPVLPSSGLPLPTPDARVCASTGCPDHLPLPGTLPDPDW-VTL 475
OY 254 DEDFVLYALHSHIGSEMFAAPMRCACAGVHMLFYVRAGVSRAMLRLFLAMEGRHME 313
DB 476 ECGFVLMALISPSHIGADLVAAPIARDDQVHLCVNRGSISSALRLFLAMEGRSHFS 535
OY 314 YECPLYVYVVAARFLERKDGKGFANVDSGLMSEAVNGQVPTFMVNSGVCEPPPSMK 373
DB 536 LCGFOLGYAAARAFRLERPLTRGVLYVDGEVYGPLOAMHPIGTLTGTG---PCC- 590
OY 374 PQMPPEEP 383
DB 591 -----PGRPEP 595

RESULT 11

US-60-453-050-7837
Sequence 7837, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARCILL, Michele
APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO01457
CURRENT APPLICATION NUMBER: US/60/453.050
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7837
LENGTH: 595
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-050-7837

Query Match 44.0% Score 886.5; DB 7; Length 595;
Best Local Similarity 39.8%; Pred. No. 1.5e-74;
Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

OY 10 VLRPCRVLLVLPNGCKKALQLEFRSHVOPLLAEAFISFTLMLTERNNHARELVRSSEEL 69
DB 117 LRPRLRLLLVNPFGGGLAMQCKNHNVLPMISEAGLSFNLQTERNNHARELVQGLSL 176
OY 70 GRMDALVYVSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 129
DB 177 SEMDGIYVYSGDGLHNEVYVNGLMERPMETAIOKPLCSLPAGSGNALASLNHYAGYEOV 236
OY 130 TNEDLTNGTLLCRRLSPMNLISLHTASGLRFLSVSLAMGFADVDLESEKRYRLGE 189
DB 237 LGIDLNLNCSLLCCGGCHPLDLISVTLASGSCFSLVAMGFVSDVDIOSEKFRALGS 296

: PRIOR APPLICATION NUMBER: 60/237,768
 : PRIOR FILING DATE: 2000-10-03
 : PRIOR APPLICATION NUMBER: 60/231,498
 : PRIOR FILING DATE: 2000-09-08
 : NUMBER OF SEQ ID NOS: 207012
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 9811
 : LENGTH: 312
 : TYPE: PRT
 : ORGANISM: Human
 : US-09-949-016-9811

Query Match 81.4%; Score 1642; DB 5; Length 312;
 Best Local Similarity 100.0%; Pred. No. 4,4e-146;
 Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALVYMSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 132
 DB 1 DALVYMSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
 QY 133 DLITNCTLLCRRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 192
 DB 61 DLITNCTLLCRRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 120
 QY 193 TLGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 252
 DB 121 TLGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
 QY 253 PDEDVLYLALHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 312
 DB 181 PDEDVLYLALHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 240
 QY 313 EXECPLVVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 372
 DB 241 EXECPLVVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 300
 QY 373 KPOMPPEEPL 384
 DB 301 KPOMPPEEPL 312

RESULT 7
 US-60-453-135-13597
 : Sequence 13597, Application US/60453135
 : GENERAL INFORMATION:
 : APPLICANT: CARGILL, Michele
 : APPLICANT: IAKOUBOVA, Olga
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : FILE REFERENCE: C1001456
 : CURRENT APPLICATION NUMBER: US/60/453,135
 : CURRENT FILING DATE: 2003-03-10
 : NUMBER OF SEQ ID NOS: 82762
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 13597
 : LENGTH: 307
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-60-453-135-13597

Query Match 80.4%; Score 1620; DB 7; Length 307;
 Best Local Similarity 100.0%; Pred. No. 5e-144;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 MSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 137
 DB 1 MSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
 QY 138 CTLTLCRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 197
 DB 61 CTLTLCRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 120
 QY 198 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 257

DB 121 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
 QY 258 VLVALLHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 317
 DB 181 VLVALLHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 240
 QY 318 YLVYVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 377
 DB 241 YLVYVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 300
 QY 378 PPEEPL 384
 DB 301 PPEEPL 307

RESULT 8
 US-60-453-050-13597
 : Sequence 13597, Application US/60453050
 : GENERAL INFORMATION:
 : APPLICANT: CARGILL, Michele
 : APPLICANT: LUKE, May
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : FILE REFERENCE: C1001457
 : CURRENT APPLICATION NUMBER: US/60/453,050
 : CURRENT FILING DATE: 2003-03-10
 : NUMBER OF SEQ ID NOS: 82762
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 13597
 : LENGTH: 307
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-60-453-050-13597

Query Match 80.4%; Score 1620; DB 7; Length 307;
 Best Local Similarity 100.0%; Pred. No. 5e-144;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 MSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 137
 DB 1 MSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALAASLNHYAGYEQVTNE 60
 QY 138 CTLTLCRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 197
 DB 61 CTLTLCRLSPMNLSTHTASGLRLEFSVLSLAWGFIADVDLESEKTRRIGEMRF 120
 QY 198 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 257
 DB 121 LRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVPLEEPPVSHMTVV 180
 QY 258 VLVALLHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 317
 DB 181 VLVALLHSHLSGSEMFAPMGRCAAGVNHLYFVAGVSRAMLRLFLAMEKGRHM 240
 QY 318 YLVYVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 377
 DB 241 YLVYVPVVAFRLEPKDGKGFVAVDGLMVSFAVGOVHPNFMVSGCVPPEPSM 300
 QY 378 PPEEPL 384
 DB 301 PPEEPL 307

RESULT 9
 US-60-452-680-12558
 : Sequence 12558, Application US/60452680
 : GENERAL INFORMATION:
 : APPLICANT: CARGILL, Michele
 : APPLICANT: GRUBE, Andrew
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : FILE REFERENCE: C1001450

Mon May.19 10:50:39 2003

us-09-937-060a-5.i.xapn

Page 3

Best Local Similarity 99.2%; Pred. No. 2, 1e-180;
Matches 361; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 MPBAGPRVLPVPRVLYVLLNPNPGSGKALQFRSHVOPLLAEAEISFTLMLERRNHA 60
Db 1 MDPAGGRVGLVPRVLYVLLNPNPGSGKALQFRSHVOPLLAEAEISFTLMLERRNHA 60
OY 61 RELVSEELGRMDALVWVGDLGMEVYNGIMERPDMETAIOKPLCSLPAGSGNALAST 120
Db 61 RELVSEELGRMDALVWVGDLGMEVYNGIMERPDMETAIOKPLCSLPAGSGNALAST 120
OY 121 NHVAGTEOYTNEDLLTNCTLLCRLLSPMNLSTHTASGLRFLSVLSIANGFTADVDLE 180
Db 121 NHVAGTEOYTNEDLLTNCTLLCRLLSPMNLSTHTASGLRFLSVLSIANGFTADVDLE 180
OY 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRILAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
Db 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRILAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEPVPVSHMTVPDEDFVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYRAGVSRAML 300
Db 241 LEPVPVSHMTVPDEDFVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYRAGVSRAML 300
OY 301 RLFLAMEKGRHMEYECPTLVYVVAFLRLEPKDKGVFAVDEGLMVAAGOVHPNTEFW 360
Db 301 RLFLAMEKGRHMEYECPTLVYVVAFLRLEPKDKGVFAVDEGLMVAAGOVHPNTEFW 360
OY 361 MWSGCVPEPPPSMKPOOMPPEEPL 384
Db 361 MWSGCVPEPPPSMKPOOMPPEEPL 384
```

RESULT 4

```
US-10-348-052-21
; Sequence 21, Application US/10348052
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION
; FILE REFERENCE: 200116.405
; CURRENT APPLICATION NUMBER: US/10/348.052
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-348-052-21
```

Query Match 95.4%; Score 1923; DB 6; Length 368;

Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 17 VLVLLNPNPGSGKALQFRSHVOPLLAEAEISFTLMLERRNHAELVRSSEELGRMDALV 76
Db 1 VLVLLNPNPGSGKALQFRSHVOPLLAEAEISFTLMLERRNHAELVRSSEELGRMDALV 76
OY 77 VWSGDLGMEVYNGIMERPDMETAIOKPLCSLPAGSGNALASTNHVAGTEOYTNEDLLT 136
Db 77 VWSGDLGMEVYNGIMERPDMETAIOKPLCSLPAGSGNALASTNHVAGTEOYTNEDLLT 136
OY 137 NCTLLCRLLSPMNLSTHTASGLRFLSVLSIANGFTADVDLESEKYRRLGEMRFTLGT 196
Db 137 NCTLLCRLLSPMNLSTHTASGLRFLSVLSIANGFTADVDLESEKYRRLGEMRFTLGT 196
OY 197 FLRLAALRTYGRILAYLPVGRVSKTPASPVVVOGQPVDAHLVLEPVPVSHMTVPDEDE 256
Db 197 FLRLAALRTYGRILAYLPVGRVSKTPASPVVVOGQPVDAHLVLEPVPVSHMTVPDEDE 256
OY 257 FVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYRAGVSRAMLRLFLAMEKGRHMEYEC 316
Db 257 FVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYRAGVSRAMLRLFLAMEKGRHMEYEC 316
```

```
Db 241 FVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYRAGVSRAMLRLFLAMEKGRHMEYEC 300
OY 317 PYLVYVVAFLRLEPKDKGVFAVDEGLMVAAGOVHPNTEFWMWSGCVPEPPPSMKPOO 376
Db 301 PYLVYVVAFLRLEPKDKGVFAVDEGLMVAAGOVHPNTEFWMWSGCVPEPPPSMKPOO 360
OY 377 MPEPEEPL 384
Db 361 MPEPEEPL 368
```

RESULT 5

```
US-10-264-237-2585
; Sequence 2585, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P433p1
; CURRENT APPLICATION NUMBER: US/10/264.237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/0501/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO: 2585
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2585
```

Query Match 86.9%; Score 1752; DB 6; Length 333;

Best Local Similarity 100.0%; Pred. No. 2, 1e-156;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 52 MLTERRNHARELVSEELGRMDALVWVGDLGMEVYNGIMERPDMETAIOKPLCSLPAG 111
Db 1 MLTERRNHARELVSEELGRMDALVWVGDLGMEVYNGIMERPDMETAIOKPLCSLPAG 60
OY 112 SGNALAASLNHYAGYBOYTNEDLLTNCTLLCRLLSPMNLSTHTASGLRFLSVLSIANG 171
Db 61 SGNALAASLNHYAGYBOYTNEDLLTNCTLLCRLLSPMNLSTHTASGLRFLSVLSIANG 120
OY 172 GFADVDLESEKYRRLGEMRFTLGTFLRLAALRTYGRILAYLPVGRVSKTPASPVVVOO 231
Db 121 GFADVDLESEKYRRLGEMRFTLGTFLRLAALRTYGRILAYLPVGRVSKTPASPVVVOO 180
OY 232 GPVDAHLVLEPVPVSHMTVPDEDFVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYR 291
Db 181 GPVDAHLVLEPVPVSHMTVPDEDFVLVLLSHLSGSEMPAAMPGRCAAGVMHLFVYR 240
OY 292 ACVSRAMLRLFLAMEKGRHMEYECPTLVYVVAFLRLEPKDKGVFAVDEGLMVAAGV 351
Db 241 ACVSRAMLRLFLAMEKGRHMEYECPTLVYVVAFLRLEPKDKGVFAVDEGLMVAAGV 300
OY 352 GOVHPNTEFWMWSGCVPEPPPSMKPOOMPPEEPL 384
Db 301 GOVHPNTEFWMWSGCVPEPPPSMKPOOMPPEEPL 333
```

RESULT 6

```
US-09-949-016-9811
; Sequence 9811, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

ORGANISM: Homo sapiens
PCT-US03-02588-26

Query Match 100.0%; Score 2016; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.7e-181;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGPGGVLPKPCRVLYLNPBGKGAQLOFRSHVQPLAEAEISFTMLTERRNA 60
DB 1 MDPAGPGGVLPKPCRVLYLNPBGKGAQLOFRSHVQPLAEAEISFTMLTERRNA 60

QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDWETAIOKPLCSIPASGNALASL 120
DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDWETAIOKPLCSIPASGNALASL 120

QY 121 NHYAGEQVNTMEDLLTNCITLLCRRLSPMNLISHTASGRLFSVLSLAWCFIADVLE 180
DB 121 NHYAGEQVNTMEDLLTNCITLLCRRLSPMNLISHTASGRLFSVLSLAWCFIADVLE 180

QY 181 SEKYRLEGMRFTLGTFLRLAALRTYGRRLAYLPYGRVSGKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRLEGMRFTLGTFLRLAALRTYGRRLAYLPYGRVSGKTPASPVVVOGSPVDAHLP 240

QY 241 LEEPVSHMTVPDEDFVLVYALHSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300
DB 241 LEEPVSHMTVPDEDFVLVYALHSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300

QY 301 RLFLAMEGRHMEYCPPLVYVVAFRLEPKDKGFAVDGELMVSEAVOGVHPNFW 360
DB 301 RLFLAMEGRHMEYCPPLVYVVAFRLEPKDKGFAVDGELMVSEAVOGVHPNFW 360

QY 361 MVSGCVPPEPSMKPOOMPPEEPL 384
DB 361 MVSGCVPPEPSMKPOOMPPEEPL 384

RESULT 2
US-10-354-358-26
Sequence 26, Application US/10354358
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-020P1NOMIUM
CURRENT APPLICATION NUMBER: US/10/354,358
CURRENT FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 26
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-10-354-358-26

Query Match 100.0%; Score 2016; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.7e-181;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGPGGVLPKPCRVLYLNPBGKGAQLOFRSHVQPLAEAEISFTMLTERRNA 60
DB 1 MDPAGPGGVLPKPCRVLYLNPBGKGAQLOFRSHVQPLAEAEISFTMLTERRNA 60

QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDWETAIOKPLCSIPASGNALASL 120
DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDWETAIOKPLCSIPASGNALASL 120

QY 121 NHYAGEQVNTMEDLLTNCITLLCRRLSPMNLISHTASGRLFSVLSLAWCFIADVLE 180
DB 121 NHYAGEQVNTMEDLLTNCITLLCRRLSPMNLISHTASGRLFSVLSLAWCFIADVLE 180

QY 181 SEKYRLEGMRFTLGTFLRLAALRTYGRRLAYLPYGRVSGKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRLEGMRFTLGTFLRLAALRTYGRRLAYLPYGRVSGKTPASPVVVOGSPVDAHLP 240

QY 241 LEEPVSHMTVPDEDFVLVYALHSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300
DB 241 LEEPVSHMTVPDEDFVLVYALHSHLSEMFAPMGRCAAGVNHLEFYRAGVSRAML 300

QY 301 RLFLAMEGRHMEYCPPLVYVVAFRLEPKDKGFAVDGELMVSEAVOGVHPNFW 360
DB 301 RLFLAMEGRHMEYCPPLVYVVAFRLEPKDKGFAVDGELMVSEAVOGVHPNFW 360

QY 361 MVSGCVPPEPSMKPOOMPPEEPL 384
DB 361 MVSGCVPPEPSMKPOOMPPEEPL 384

RESULT 3
US-09-949-016-7026
Sequence 7026, Application US/09949016
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7026
LENGTH: 384
TYPE: PRT
ORGANISM: Human
US-09-949-016-7026

Query Match 99.6%; Score 2008; DB 5; Length 384;

Mon May 19 10:50:39 2003

us-09-937-060a-5.rapn

Page 1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:48:53 ; Search time 37 Seconds
(without alignments)
1774.784 Million cells updates/sec

Title: US-09-937-060a-5
Perfect score: 2016
Sequence: 1 MDPAGCPKGVLPKPCRVLYL.....CVPFPMKPKQMPPEPEPL 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 787003 seqs, 171007862 residues

Total number of hits satisfying chosen parameters: 787003

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Pending Patents AA-New: *
1: /cgn2_6/pdata/2/paa/US02_NEW.COMB.pep: *
2: /cgn2_6/pdata/2/paa/US07_NEW.COMB.pep: *
3: /cgn2_6/pdata/2/paa/US07_NEW.COMB.pep: *
4: /cgn2_6/pdata/2/paa/US08_NEW.COMB.pep: *
5: /cgn2_6/pdata/2/paa/US09_NEW.COMB.pep: *
6: /cgn2_6/pdata/2/paa/US10_NEW.COMB.pep: *
7: /cgn2_6/pdata/2/paa/US60_NEW.COMB.pep: *
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	2016	100.0	384	1	PCT-US03-02588-26
2	2016	100.0	384	6	US-10-354-358-26
3	2008	99.6	384	5	US-09-949-016-7026
4	1923	95.4	368	6	US-10-348-052-21
5	1752	86.9	333	6	US-10-264-237-2585
6	1642	81.4	312	5	US-09-949-016-9811
7	1620	80.4	307	7	US-60-453-135-13597
8	1620	80.4	307	7	US-60-452-680-12558
9	886.5	44.0	595	7	US-60-453-135-7837
10	886.5	44.0	595	7	US-60-453-050-7837
11	886.5	44.0	595	7	US-60-453-050-7837
12	886.5	44.0	595	7	US-60-453-050-7837
13	886.5	44.0	595	7	US-60-453-050-7837
14	886.5	44.0	595	7	US-60-453-050-7837
15	886.5	44.0	595	7	US-60-453-050-7837
16	886.5	44.0	595	7	US-60-453-050-7837
17	886.5	44.0	595	7	US-60-453-050-7837
18	886.5	44.0	595	7	US-60-453-050-7837
19	886.5	44.0	595	7	US-60-453-050-7837
20	886.5	44.0	595	7	US-60-453-050-7837
21	562	27.9	490	6	US-10-348-052-28
22	521.5	25.9	524	6	US-10-348-052-29
23	521.5	25.9	524	6	US-10-348-052-29
24	363.5	18.0	349	7	US-60-453-135-7838
25	363.5	18.0	349	7	US-60-453-135-7838
26	363.5	18.0	349	7	US-60-453-135-7838

27	264.5	13.1	537	1	PCT-US02-31373-40	Sequence 40, Appl
28	213.5	10.6	422	6	US-10-111-896-5	Sequence 5, Appl
29	207.5	10.3	420	6	US-10-111-896-3	Sequence 3, Appl
30	202	10.0	421	6	US-10-111-896-7	Sequence 7, Appl
31	155	7.9	182	6	US-10-218-140-4294	Sequence 4294, Appl
32	144.5	7.2	350	6	US-10-282-122A-53800	Sequence 53800, A
33	144.5	7.0	322	6	US-10-156-761-10552	Sequence 10552, A
34	136	6.7	319	6	US-10-156-761-9811	Sequence 9811, Ap
35	132.5	6.6	304	5	US-09-134-000C-5366	Sequence 5366, Ap
36	132.5	6.6	304	5	US-09-134-000C-5366	Sequence 5366, Ap
37	129.5	6.4	325	6	US-10-417-884-6786	Sequence 6786, Ap
38	126.5	6.3	305	6	US-10-282-122A-61429	Sequence 61429, A
39	125	6.2	340	6	US-10-282-122A-74466	Sequence 74466, A
40	124	6.2	345	5	US-09-134-000C-3813	Sequence 3813, Ap
41	124	6.2	345	5	US-09-134-000C-3813	Sequence 3813, Ap
42	124	6.2	362	6	US-10-282-122A-42541	Sequence 42541, A
43	120.5	6.0	297	5	US-09-134-000C-3616	Sequence 3616, Ap
44	120.5	6.0	297	5	US-09-134-000C-3616	Sequence 3616, Ap
45	120.5	6.0	304	1	PCT-US02-36123-4090	Sequence 4090, Ap

ALIGNMENTS

RESULT 1
PCT-US03-02588-26
Sequence 26, Application PC/TUS0302588
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesson, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16314, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-020
CURRENT APPLICATION NUMBER: PCT/US03/02588
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
REMAINING Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 384
TYPE: PRT


```

Db 61 VMSGDGLMEVYVNGIMERPDMETAIQKPLCSLPGSGNALASLHNYAGYEOVNTEDLLT 120
QY 137 NCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLWGFIDVDESEKRYRIGEMRFLGT 136
Db 121 NCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLWGFIDVDESEKRYRIGEMRFLGT 180
QY 197 FLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQGVPAHLVPLEEVPVSHWTVPDED 256
Db 181 FLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQGVPAHLVPLEEVPVSHWTVPDED 240
QY 257 FVLVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYRAGVSRAHLFLFLAMEKGRHMETEC 316
Db 241 FVLVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYRAGVSRAHLFLFLAMEKGRHMETEC 300
QY 317 FVLVYVYVVAERFLERKDGKGFVAVDGLMVSNAVQGVHPNTFMVNSGCVPEPPSMKPOQ 376
Db 301 FVLVYVYVVAERFLERKDGKGFVAVDGLMVSNAVQGVHPNTFMVNSGCVPEPPSMKPOQ 360
QY 377 MPEPEEPL 384
Db 361 MPEPEEPL 368

RESULT 14
PCT-US01-16450-2585
: Sequence 2585, Application PC/TUS0116450
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA131PCT
: CURRENT APPLICATION NUMBER: PCT/US01/16450
: CURRENT FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: 60/205,515
: PRIOR FILING DATE: 2000-05-19
: NUMBER OF SEQ ID NOS: 2820
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2585
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US01-16450-2585

```

```

Query Match 86.9%; Score 1752; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 8.7e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 52 MLTERRNHARELVSEELGRMDALVVMGSGDGLMHEVYVNGIMERPDMETAIQKPLCSLPAG 111
Db 1 MLTERRNHARELVSEELGRMDALVVMGSGDGLMHEVYVNGIMERPDMETAIQKPLCSLPAG 60
QY 112 SGNALASLHNYAGYEOVNTEDLLTCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLAW 171
Db 61 SGNALASLHNYAGYEOVNTEDLLTCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLAW 120
QY 172 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQ 231
Db 121 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQ 180
QY 232 GPVDALHVLPEEVPVSHWTVPDEDVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYR 291
Db 181 GPVDALHVLPEEVPVSHWTVPDEDVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYR 240
QY 292 AGVSRAMLLRFLFLAMEKGRHMETECRYLVYVVAERFLERKDGKGFVAVDGLMVSNAVQ 351
Db 241 AGVSRAMLLRFLFLAMEKGRHMETECRYLVYVVAERFLERKDGKGFVAVDGLMVSNAVQ 300
QY 352 GOVHPNTFMVNSGCVPEPPSMKPOQMPPEEPL 384
Db 301 GOVHPNTFMVNSGCVPEPPSMKPOQMPPEEPL 333

```

RESULT 15

```

PCT-US01-16450A-2585
: Sequence 2585, Application PC/TUS0116450A
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA131PCT
: CURRENT APPLICATION NUMBER: PCT/US01/16450A
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 60/205,515
: PRIOR FILING DATE: 2000-05-19
: NUMBER OF SEQ ID NOS: 2820
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2585
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US01-16450A-2585

```

```

Query Match 86.9%; Score 1752; DB 1; Length 333;
Best Local Similarity 100.0%; Pred. No. 8.7e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 52 MLTERRNHARELVSEELGRMDALVVMGSGDGLMHEVYVNGIMERPDMETAIQKPLCSLPAG 111
Db 1 MLTERRNHARELVSEELGRMDALVVMGSGDGLMHEVYVNGIMERPDMETAIQKPLCSLPAG 60
QY 112 SGNALASLHNYAGYEOVNTEDLLTCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLAW 171
Db 61 SGNALASLHNYAGYEOVNTEDLLTCTLLCRRLSPNNLSLTFTASGLRFLFSVLSLAW 120
QY 172 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQ 231
Db 121 GFLADVDESEKRYRIGEMRFLGTFLRLAALRTYRGRLAYLPVGVGSKTPASPVYVQ 180
QY 232 GPVDALHVLPEEVPVSHWTVPDEDVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYR 291
Db 181 GPVDALHVLPEEVPVSHWTVPDEDVYVLAALHSHLGSFMPAARPMGRCAAGVNHLEFVYR 240
QY 292 AGVSRAMLLRFLFLAMEKGRHMETECRYLVYVVAERFLERKDGKGFVAVDGLMVSNAVQ 351
Db 241 AGVSRAMLLRFLFLAMEKGRHMETECRYLVYVVAERFLERKDGKGFVAVDGLMVSNAVQ 300
QY 352 GOVHPNTFMVNSGCVPEPPSMKPOQMPPEEPL 384
Db 301 GOVHPNTFMVNSGCVPEPPSMKPOQMPPEEPL 333

```

Search completed: May 9, 2003, 16:52:50
Job time : 152 secs

NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent version 3.1
SEQ ID NO 3
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(384)
OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspond-
OTHER INFORMATION: nating to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
PUBLICATION INFORMATION: nbank sequence Accession Number AAF73423.
AUTHORS: Nava et al.
TITLE: Functional characterization of human sphingosine kinase-1.
JOURNAL: FEBS Lett.
VOLUME: 473
ISSUE: 1
PAGES: 81-84
DATE: 2000
DATABASE ACCESSION NUMBER: AAF73423
DATABASE ENTRY DATE: 2000-06-01
RELEVANT RESIDUES: (1)..(384)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AAF73423
DATABASE ENTRY DATE: 2000-06-01
RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match 99.6%; Score 2008; DB 21; Length 384;
Best Local Similarity 99.2%; Pred. No. 8.5e-203;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCVVLVILNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
DB 1 MDPAGPRGVLPKPCVVLVILNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
QY 61 RELVRSSEELGRMDALVMSGDDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVRSSEELGRMDALVMSGDDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGYEQVNTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGYEQVNTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKRYRLGEMRFTLGFTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVVDHILVP 240
DB 181 SEKRYRLGEMRFTLGFTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVVDHILVP 240
QY 241 LEEPVPSHMTVVPDEDFVLVLAALSHSLGSEMFAPMGRCAAGVMHLFVVRAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDFVLVLAALSHSLGSEMFAPMGRCAAGVMHLFVVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECFYLTVYVVAFAFLRLEPKDKGVFAVDGLMWSSEAVOGVHPNYFM 360
DB 301 RLFLAMEKGRHMEYECFYLTVYVVAFAFLRLEPKDKGVFAVDGLMWSSEAVOGVHPNYFM 360
QY 361 MVSQGVPEPPSMKPOOMPPEEPL 384
DB 361 MVSQGVPEPPSMKPOOMPPEEPL 384

RESULT 12
US-09-784-810A-2

Sequence 2, Application US/09784810A
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
TITLE OF INVENTION: SAME
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360

PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-784-810A-2

Query Match 97.3%; Score 1962; DB 21; Length 384;
Best Local Similarity 97.7%; Pred. No. 6.3e-198;
Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCVVLVILNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
DB 1 MDPAGPRGVLPKPCVVLVILNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNA 60
QY 61 RELVRSSEELGRMDALVMSGDDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVRSSEELGRMDALVMSGDDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
QY 121 NHYAGYEQVNTNEDLTNCTLLCRLLSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
DB 121 NHYAGYEQVNTNEDLTNCTLLCRPLVSPMNLISHTASGLRFSVLSLAWGFADVDLE 180
QY 181 SEKRYRLGEMRFTLGFTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVVDHILVP 240
DB 181 SEKRYRLGEMRFTLGFTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOOGVVDHILVP 240
QY 241 LEEPVPSHMTVVPDEDFVLVLAALSHSLGSEMFAPMGRCAAGVMHLFVVRAGVSRAML 300
DB 241 LEEQVSHMQVVPDEDFVLVLAALSHSLGSEMFAPMGRCAAGVMHLFVVRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECFYLTVYVVAFAFLRLEPKDKGVFAVDGLMWSSEAVOGVHPNYFM 360
DB 301 RLFLAMEKGRHMEYECFYLTVYVVAFAFLRLEPKDKGVFAVDGLMWSSEAVOGVHPNYFM 360
QY 361 MVSQGVPEPPSMKPOOMPPEEPL 384
DB 361 MVSQGVPEPPSMKPOOMPPEEPL 384

RESULT 13
US-10-053-510-21

Sequence 21, Application US/10053510
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
FILE REFERENCE: 200116.402C2
CURRENT APPLICATION NUMBER: US/10/053,510
CURRENT FILING DATE: 2002-01-17
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 368
TYPE: PRT
ORGANISM: Homo sapiens
US-10-053-510-21

Query Match 95.4%; Score 1923; DB 24; Length 368;
Best Local Similarity 100.0%; Pred. No. 8e-194;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 VLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNAHRELVSSEELGRMDALV 76
DB 1 VLVLNPRGKGKALQIFRSHVQPLAEAFISFTMLTERRNAHRELVSSEELGRMDALV 60
QY 77 VMSGDDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGYEQVNTNEDLT 136

Mon May 19 10:50:39 2003

us-09-937-060a-5. rapm

Page 5

```

OY 241 LEEPPVSHMTVPDEDFVLTALILSHSGSEMFAPRGCAACVNHLEFVAGVSRAML 300
Db 241 LEEPPVSHMTVPDEDFVLTALILSHSGSEMFAPRGCAACVNHLEFVAGVSRAML 300
OY 301 RLFLAMEGRHMECECYLVYVYVVAFLREKDDKGVFAVDGLAMVSEAVOGVHPNPFV 360
Db 301 RLFLAMEGRHMECECYLVYVYVVAFLREKDDKGVFAVDGLAMVSEAVOGVHPNPFV 360
OY 361 MVSGCEPPPSMKQOMPPEPL 384
Db 361 MVSGCEPPPSMKQOMPPEPL 384

RESULT 9
US-09-488-725a-3325
? Sequence 3325, Application US/09488725A
? GENERAL INFORMATION:
? APPLICATION: Hyseq Inc
? TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
? FILE REFERENCE: 784FLPCT
? CURRENT APPLICATION NUMBER: US/09/488, 725A
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: US/09/488, 725
? PRIOR FILING DATE: 2000-01-21
? PRIOR APPLICATION NUMBER: US09/552, 317
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: US09/598, 042
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: US09/620, 312
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: US09/653, 450
? PRIOR FILING DATE: 2000-08-31
? PRIOR APPLICATION NUMBER: US09/662, 191
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: US09/693, 036
? PRIOR FILING DATE: 2000-10-19
? PRIOR APPLICATION NUMBER: US09/727, 344
? PRIOR FILING DATE: 2000-11-29
? NUMBER OF SEQ ID NOS: 7144
? SOFTWARE: pL_FL_genes_b Versions 1.0
? SEQ ID NO 3325
? LENGTH: 384
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-488-725a-3325

Query Match 99.8%; Score 2011; DB 18; Length 384;
Best Local Similarity 99.7%; Pred. No. 4,1e-203;
Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MDAGCGPGCVLPACRCRYVLINLRPGKGKAAOLFRSHVOPILAEAFISTMLTERRNA 60
Db 1 MDAGCGPGCVLPACRCRYVLINLRPGKGKAAOLFRSHVOPILAEAFISTMLTERRNA 60
OY 61 RELVSESEELGEMDALVYMSGDGLMEHYVNGLMERPMDETAIOKPLCLSPAGSGNALAASL 120
Db 61 RELVSESEELGEMDALVYMSGDGLMEHYVNGLMERPMDETAIOKPLCLSPAGSGNALAASL 120
OY 121 NHYACGEOVTMEDLLTTCCTLLCRLLSPMNLSLHFAASGLRLEFVLSLAWGFIADNDLE 180
Db 121 NHYACGEOVTMEDLLTTCCTLLCRLLSPMNLSLHFAASGLRLEFVLSLAWGFIADNDLE 180
OY 181 SEKYRRLGEMRFTLGFTLRILAALIRYRKGHLALPYGVGSKTASAVYVYVQGDPAHILVP 240
Db 181 SEKYRRLGEMRFTLGFTLRILAALIRYRKGHLALPYGVGSKTASAVYVYVQGDPAHILVP 240
OY 241 LEEPPVSHMTVPDEDFVLTALILSHSGSEMFAPRGCAACVNHLEFVAGVSRAML 300
Db 241 LEEPPVSHMTVPDEDFVLTALILSHSGSEMFAPRGCAACVNHLEFVAGVSRAML 300
OY 301 RLFLAMEGRHMECECYLVYVYVVAFLREKDDKGVFAVDGLAMVSEAVOGVHPNPFV 360
Db 301 RLFLAMEGRHMECECYLVYVYVVAFLREKDDKGVFAVDGLAMVSEAVOGVHPNPFV 360

```

```

Oy      361 MVSQVEPPPSMKPQQMPPEEPL 364
          |||||||
Db      361 MVSQVEPPPSMKPQQMPPEEPL 364

RESULT 10
PCT-US02-24623-10
? Sequence 10, Application PC/TUS0224623
? GENERAL INFORMATION:
? APPLICANT: EXELTIS, INC.
? TITLE OF INVENTION: SPIKS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
? FILE REFERENCE: EX02-104
? CURRENT APPLICATION NUMBER: PCT/US02/24623
? PRIOR FILING DATE: 2002-08-02
? PRIOR APPLICATION NUMBER: US 60/310,362
? PRIOR FILING DATE: 2001-08-06
? PRIOR APPLICATION NUMBER: US 60/357,501
? PRIOR FILING DATE: 2002-02-15
? NUMBER OF SEQ ID NOS: 13
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 10
? LENGTH: 384
? TYPE: PRT
? ORGANISM: Homo sapiens
PCT-US02-24623-10

Query Match           99.6%; Score 2008; DB 1; Length 384;
Best Local Similarity 99.2%; Pred. No. 8, Se=203;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY       1 MDPAAGCGVLPAPCCFVLVLNPRGKGKALQLFRSHVOPLLAEISETLMTERRHA 60
DB       1 MDPAGCGVLPAPCCFVLVLNPRGKGKALQLFRSHVOPLLAEISETLMTERRHA 60

OY       61 RELVASERLRMADLVMSGDGLMHEVVNGIMERPDMETAIQKPLCLSPAGSGNALAASL 120
DB       61 RELVASERLRMADLVMSGDGLMHEVVNGIMERPDMETAIQKPLCLSPAGSGNALAASL 120

OY       121 NHAYAGEQVTNEDLLNCTLLRCRLLSPNNLLSHTASGLRLFSVYSLAMGFADVPLE 180
DB       121 NHAYAGEQVTNEDLLNCTLLRCRLLSPNNLLSHTASGLRLFSVYSLAMGFADVPLE 180

OY       181 SKRYARLGEMRFLLFTFLRLATLTGRGLATLYPCVRGSKTPASPVPYVOOGGVNAHLHP 240
DB       181 SKRYARLGEMRFLLFTFLRLATLTGRGLATLYPCVRGSKTPASPVPYVOOGGVNAHLHP 240

OY       241 LEEPYPMWTVVDEDFVYLVALHSHLSGESEFAAPMGCAACAAWHLFYTRACYSRAHLL 300
DB       241 LEEPYPMWTVVDEDFVYLVALHSHLSGESEFAAPMGCAACAAWHLFYTRACYSRAHLL 300

OY       301 RLFLAMEGRHMEECFYLIVPVVAFLERLPDKDGKVFAYVDDELVAEAVGGVPHNFVW 360
DB       301 RLFLAMEGRHMEECFYLIVPVVAFLERLPDKDGKVFAYVDDELVAEAVGGVPHNFVW 360

OY       361 MVSQVEPPPSMKPQQMPPEEPL 364
          |||||||
DB       361 MVSQVEPPPSMKPQQMPPEEPL 364

RESULT 11
US-09-796-487-3
? Sequence 3, Application US/09796487
? GENERAL INFORMATION:
? APPLICANT: Spiegel, Sarah
? TITLE OF INVENTION: Splicing Kinase, Cloning, Expression and Methods of Use
? FILE REFERENCE: 0732001aa (203395/-0001)
? CURRENT APPLICATION NUMBER: US/09/796,487
? CURRENT FILING DATE: 2001-03-02
? PRIOR APPLICATION NUMBER: US 60/186,532
? PRIOR FILING DATE: 2000-03-03
? PRIOR APPLICATION NUMBER: US 09/530,868
? PRIOR FILING DATE: 2000-05-05
```

```
;; TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
;; FILE REFERENCE: PF-0683 P
;; CURRENT APPLICATION NUMBER: US/60/125,593
;; CURRENT FILING DATE: 1999-03-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PERL Program
;; SEQ ID NO 5
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: 2415617
US-60-125-593-5

Query Match          100.0%; Score 2016; DB 27; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
DB 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
OY 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
DB 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPVSHTVVPDDEDFVLYLALLSHSGSEMFAPMGKCAAGVNHLEFVRAGVSRAAML 300
DB 241 LEEPVSHTVVPDDEDFVLYLALLSHSGSEMFAPMGKCAAGVNHLEFVRAGVSRAAML 300
OY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNFTW 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNFTW 360
OY 361 MVSQCEPPEPPSKPQOMPEPEPL 384
DB 361 MVSQCEPPEPPSKPQOMPEPEPL 384

RESULT 7
US-09-959-897-2
;; Sequence 2, Application US/09959897
;; GENERAL INFORMATION:
;; APPLICANT: PITSON, Stuart M
;; APPLICANT: Brian, MATTEBERG W
;; APPLICANT: Pu, XIA
;; APPLICANT: Richard, D'ANDREA J
;; APPLICANT: Jennifer, BAMBLE R
;; APPLICANT: Mathew, VADAS A
;; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
;; FILE REFERENCE: PITSON-1
;; CURRENT APPLICATION NUMBER: US/09/959, 897
;; CURRENT FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: PCT/AU00/00457
;; PRIOR FILING DATE: 2000-05-12
;; PRIOR APPLICATION NUMBER: AU PO 0339
;; PRIOR FILING DATE: 1999-05-13
;; PRIOR APPLICATION NUMBER: AU PO 1504
;; PRIOR FILING DATE: 1999-07-08
;; NUMBER OF SEQ ID NOS: 56
;; SOFTWARE: PatentIn, version 3.1
;; SEQ ID NO 2
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
```

```
US-09-959-897-2

Query Match          99.9%; Score 2013; DB 23; Length 384;
Best Local Similarity 99.7%; Pred. No. 2.5e-203;
Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
DB 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
OY 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
DB 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPVSHTVVPDDEDFVLYLALLSHSGSEMFAPMGKCAAGVNHLEFVRAGVSRAAML 300
DB 241 LEEPVSHTVVPDDEDFVLYLALLSHSGSEMFAPMGKCAAGVNHLEFVRAGVSRAAML 300
OY 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNFTW 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFLRLEPKDGKGFVAVDGLMVSSEAVOGQVHPNFTW 360
OY 361 MVSQCEPPEPPSKPQOMPEPEPL 384
DB 361 MVSQCEPPEPPSKPQOMPEPEPL 384

RESULT 8
US-60-364-517-2
;; Sequence 2, Application US/60364517
;; GENERAL INFORMATION:
;; APPLICANT: Tsai, Pong-Ying
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;; FILE REFERENCE: MPT02-032P1 (M)
;; CURRENT APPLICATION NUMBER: US/60/364, 517
;; CURRENT FILING DATE: 2002-03-15
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: fastseq for windows version 4.0
;; SEQ ID NO 2
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-60-364-517-2

Query Match          99.9%; Score 2013; DB 27; Length 384;
Best Local Similarity 99.7%; Pred. No. 2.5e-203;
Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
DB 1 MDPAGPAGVLPKPCRVLYLNLNPRGKGKALQLFPSHVQPLLAEEAISTTLMTERRNA 60
OY 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYMSGDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
DB 121 NHYAGYEQVTNEDLLTNCILLCRRLSPMNLISHTASGLRFSVLSIAMGFADVDLE 180
OY 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLTGTFRLAALRTYRGLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
```

Mon May 19 10:50:39 2003

us-09-937-060a-5.ram

Page 3

OTHER INFORMATION: Incyte ID No: 2415617CD1
US-09-937-060A-5

Query Match 100.0%; Score 2016; DB 23; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60
DB 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60
QY 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120

QY 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180

QY 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240

QY 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300
DB 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300

QY 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360
DB 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360

QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 4
US-09-970-516-2
Sequence 2, Application US/09970516
GENERAL INFORMATION:
APPLICANT: Novartis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-2

Query Match 100.0%; Score 2016; DB 23; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60
DB 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60

QY 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120

QY 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180

QY 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240

QY 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300
DB 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300
QY 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360
DB 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360

QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 5
US-10-111-729-3
Sequence 3, Application US/10111729
GENERAL INFORMATION:
APPLICANT: Warner-lambert
TITLE OF INVENTION: Human sphingosine kinase gene
FILE REFERENCE: A0001986CT
CURRENT APPLICATION NUMBER: US/10/111,729
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-10-111-729-3

Query Match 100.0%; Score 2016; DB 25; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60
DB 1 MDPAGGPRGVLPKRCRYLVLLNPRGKGKALOLFHSVOPPLAEAEISFTLMTERRNA 60

QY 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYVSGDGLMEHYVNGLMERPDMETAIQKPLCSLPAGSGNALAASL 120

QY 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180
DB 121 NHYAGEQVNTNEDLLTNCITLLCRLLSPNNLLSHTASGLRFSVLSLWGFIAVDLE 180

QY 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240
DB 181 SEKYRRLGEMRFTLGTFLRLAALRTYGRRLATLPVGRVSKTPASPVVOOGPVDAHLVP 240

QY 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300
DB 241 LEEPVSHMTVPDEDFVLYVLLSHLSGSEMFAPMRCACAGVMHLEFYVAGVSRAML 300

QY 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360
DB 301 RFLAMEKGRHMEYECPLYVVPVAFRLPKDKGVFAVDGELMVSEAVOGVHPNFTM 360

QY 361 MVSQGVPEPPPSMKPOQMPPEEPL 384
DB 361 MVSQGVPEPPPSMKPOQMPPEEPL 384

RESULT 6
US-60-125-593-5
Sequence 5, Application US/60125593
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Yalda

;; PRIOR APPLICATION NUMBER: 60/159,590
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 60/183,322
;; PRIOR FILING DATE: 2000-02-17
;; NUMBER OF SEQ ID NOS: 19025
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 13997
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-629-469A-13997

Query Match 100.0%; Score 2016; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 1,2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCRYVLLNPRGKGKALQLFSSHVQPLLAEAISFTMLTERRNHA 60
DB 1 MDPAGPRGVLPKPCRYVLLNPRGKGKALQLFSSHVQPLLAEAISFTMLTERRNHA 60
QY 61 RELVSEELGRMDALVYVSGDGLMHEVYVGLMERPDMEATAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYVSGDGLMHEVYVGLMERPDMEATAIQKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEQVNTNEDLTNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFADVDLE 180
DB 121 NHYAGEQVNTNEDLTNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFADVDLE 180
QY 181 SEKYRRLEGMRRTLTGFTLRALRTYRGRLAYLPVGRVSKTPASPVVVOQGVDAHLVP 240
DB 181 SEKYRRLEGMRRTLTGFTLRALRTYRGRLAYLPVGRVSKTPASPVVVOQGVDAHLVP 240
QY 241 LEEPVSHTVVPDDEDFVYLALSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAMLL 300
DB 241 LEEPVSHTVVPDDEDFVYLALSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAMLL 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNYFM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNYFM 360
QY 361 MVSGCPEPPPSMKPQOMPPEEPL 384
DB 361 MVSGCPEPPPSMKPQOMPPEEPL 384

RESULT 2

US-09-629-469A-15394

;; Sequence 15394, Application US/09629469A
;; GENERAL INFORMATION:

;; APPLICANT: OTO, TOSHIO
;; APPLICANT: ISOGAI, TAKAO
;; APPLICANT: NISHIKAWA, TETSUO
;; APPLICANT: HAYASHI, KOJI
;; APPLICANT: SATTO, KAORU
;; APPLICANT: YAMAMOTO, JUNICHI
;; APPLICANT: ISHII, SHIZUKO
;; APPLICANT: SUGIYAMA, TOMOYASU
;; APPLICANT: WAKAMATSU, AI
;; APPLICANT: NAGAI, KEIICHI
;; APPLICANT: OTSUKI, TETSUJI
;; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;; FILE REFERENCE: 084335/0123
;; CURRENT APPLICATION NUMBER: US/09/629,469A
;; CURRENT FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: JP 1999-248036
;; PRIOR FILING DATE: 1999-07-29
;; PRIOR APPLICATION NUMBER: JP 1999-300253
;; PRIOR FILING DATE: 1999-08-27
;; PRIOR APPLICATION NUMBER: JP 2000-118776
;; PRIOR FILING DATE: 2000-01-11
;; PRIOR APPLICATION NUMBER: JP 2000-183767
;; PRIOR FILING DATE: 2000-05-02
;; PRIOR APPLICATION NUMBER: JP 2000-241899

;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/159,590
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 60/183,322
;; PRIOR FILING DATE: 2000-02-17
;; NUMBER OF SEQ ID NOS: 19025
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 15394
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-629-469A-15394

Query Match 100.0%; Score 2016; DB 20; Length 384;
Best Local Similarity 100.0%; Pred. No. 1,2e-203;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGPRGVLPKPCRYVLLNPRGKGKALQLFSSHVQPLLAEAISFTMLTERRNHA 60
DB 1 MDPAGPRGVLPKPCRYVLLNPRGKGKALQLFSSHVQPLLAEAISFTMLTERRNHA 60
QY 61 RELVSEELGRMDALVYVSGDGLMHEVYVGLMERPDMEATAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVSEELGRMDALVYVSGDGLMHEVYVGLMERPDMEATAIQKPLCSLPAGSGNALAASL 120
QY 121 NHYAGEQVNTNEDLTNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFADVDLE 180
DB 121 NHYAGEQVNTNEDLTNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFADVDLE 180
QY 181 SEKYRRLEGMRRTLTGFTLRALRTYRGRLAYLPVGRVSKTPASPVVVOQGVDAHLVP 240
DB 181 SEKYRRLEGMRRTLTGFTLRALRTYRGRLAYLPVGRVSKTPASPVVVOQGVDAHLVP 240
QY 241 LEEPVSHTVVPDDEDFVYLALSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAMLL 300
DB 241 LEEPVSHTVVPDDEDFVYLALSHLSGSEMPFAAPMGCAAGVNHLPFVRAGVSRAMLL 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNYFM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFFRLEPKDGKGFVAVDGLMVSSEAVQGVHPNYFM 360
QY 361 MVSGCPEPPPSMKPQOMPPEEPL 384
DB 361 MVSGCPEPPPSMKPQOMPPEEPL 384

RESULT 3

US-09-937-060A-5

;; Sequence 5, Application US/09937060A
;; GENERAL INFORMATION:

;; APPLICANT: INCYTE PHARMACEUTICALS, INC.
;; APPLICANT: BANDMAN, Olga
;; APPLICANT: TANG, Y. Tom
;; APPLICANT: YUE, Henry
;; APPLICANT: HILLMAN, Jennifer L.
;; APPLICANT: BAUGHN, Mariah R.
;; APPLICANT: AZIMZAI, Yalda
;; APPLICANT: LU, Dying Aina M.
;; APPLICANT: AU-YOUNG, Janice
;; TITLE OF INVENTION: REGULATORS OF INTRACELLULAR PHOSPHORYLATION
;; FILE REFERENCE: PF-0683 PCT
;; CURRENT APPLICATION NUMBER: US/09/937,060A
;; CURRENT FILING DATE: 2002-04-15
;; PRIOR APPLICATION NUMBER: 60/125,593; 60/135,049; 60/143,188
;; PRIOR FILING DATE: 1999-03-18; 1999-05-20; 1999-07-09
;; SOFTWARE: PERL Program
;; SEQ ID NO 5
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:46:23 ; Search time 149 Seconds
(without alignments)
1061.594 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPAGCPRCVLPDPCRVLVL.....CVPSPSKMPOMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 segs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2016	100.0	384	20 US-09-629-469A-13997	Sequence 13997, A
2	2016	100.0	384	20 US-09-629-469A-15394	Sequence 15394, A
3	2016	100.0	384	23 US-09-937-060A-5	Sequence 5, App11
4	2016	100.0	384	23 US-09-970-516-2	Sequence 2, App11
5	2016	100.0	384	25 US-10-111-729-3	Sequence 3, App11
6	2016	100.0	384	27 US-60-125-593-5	Sequence 5, App11

7	2013	99.9	384	23 US-09-959-897-2	Sequence 2, App11
8	2013	99.9	384	27 US-60-364-517-2	Sequence 2, App11
9	2011	99.8	384	18 US-09-488-725A-3325	Sequence 3325, Ap
10	2008	99.6	384	1 PCT-US02-24623-10	Sequence 10, App1
11	2008	99.6	384	21 US-09-796-487-3	Sequence 1, App11
12	1962	97.3	384	21 US-09-784-810A-2	Sequence 2, App11
13	1923	95.4	368	24 US-10-053-510-21	Sequence 21, App1
14	1752	86.9	333	1 PCT-US01-16450-2585	Sequence 2585, Ap
15	1752	86.9	333	1 PCT-US01-16450A-2585	Sequence 2585, Ap
16	1639.5	81.3	382	21 US-09-784-810A-4	Sequence 4, App11
17	1632.5	81.0	382	23 US-09-970-516-6	Sequence 6, App11
18	1630.5	80.9	388	1 PCT-US01-09664A-15	Sequence 15, App11
19	1630.5	80.9	388	1 US-09-796-487-2	Sequence 2, App11
20	1630.5	80.9	388	22 US-09-817-676A-15	Sequence 15, App11
21	1627.5	80.7	381	21 US-09-796-487-1	Sequence 1, App11
22	1627.5	80.7	381	21 US-09-796-487-4	Sequence 4, App11
23	1611	79.9	305	18 US-09-488-725A-6897	Sequence 6897, Ap
24	1609.5	79.8	373	21 US-09-796-487-5	Sequence 5, App11
25	1541	76.4	293	1 PCT-US01-05614-328	Sequence 328, App
26	1541	76.4	293	1 PCT-US98-11422A-313	Sequence 313, App
27	1541	76.4	293	16 US-09-205-258-328	Sequence 328, App
28	1541	76.4	293	23 US-09-933-767-328	Sequence 328, App
29	1541	76.4	293	24 US-10-004-860-328	Sequence 328, App
30	1541	76.4	293	24 US-10-023-282-328	Sequence 328, App
31	1477.5	73.3	381	27 US-60-243-468-1133	Sequence 1133, Ap
32	1077.5	53.4	260	20 US-09-641-377-255	Sequence 255, App
33	940	46.6	548	27 US-60-243-468-719	Sequence 719, App
34	938.5	46.6	204	21 US-09-796-487-9	Sequence 9, App11
35	893	44.3	617	1 PCT-US01-09664A-12	Sequence 12, App1
36	893	44.3	617	22 US-09-817-676A-12	Sequence 12, App1
37	886.5	44.0	618	1 PCT-US01-09664A-14	Sequence 14, App1
38	886.5	44.0	618	1 PCT-US02-24623-11	Sequence 11, App1
39	886.5	44.0	618	22 US-09-817-676A-14	Sequence 14, App1
40	886.5	44.0	618	22 US-09-970-516-4	Sequence 4, App11
41	886.5	44.0	618	27 US-60-388-853-4	Sequence 4, App11
42	886.5	44.0	654	20 US-09-641-377-254	Sequence 254, App
43	873.5	43.3	806	1 PCT-US01-08631-51503	Sequence 51503, A
44	833.5	41.3	638	23 US-09-969-896-3	Sequence 3, App11
45	694	34.4	146	27 US-60-178-307-3020	Sequence 3020, Ap

ALIGNMENTS

RESULT 1
US-09-629-469A-13997
Sequence 13997 Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OITA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629, 469A
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09

TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-784-810a-15

Query Match 18.6%; Score 374.5; DB 10; Length 415;
Best Local Similarity 28.6%; Pred. No. 2,9e-28;
Matches 110; Conservative 80; Mismatches 146; Indels 49; Gaps 12;

```

QY 12 PRP---CR--VLVLINPRGGKALOLFRRSHVOPDLAEAEISFTLMTERRNHARE-LVR 65
DB 20 PPPOEOCRGNILVFINNSGTGKSLFTFANTVGPRLDKSLIRYEVVYTTGPNHARNVLTMT 79
QY 66 SEELGRMDALVMSGDIIMEVNVGLMERPD-METATQKPLCSLPAGSGNALAASLNHYA 124
DB 80 KADGKFNGLVLIISGDIVPEALNGILCREDAFRIFPTLPIGIVPSGSGNGLCSV--LS 137
QY 125 GYEQVTNEDLITNCTLLCLRRLSPANLLSLHT--ASGLRFSVLSIANGFIADVDLESE 182
DB 138 KYGTMMNEKSVMERALEIATSPTKAESVALYSVKTNDQSYASFSLISGWMADIDIDSE 197
QY 183 KYRR-LGEMRPTLGTFLRLAALFTYRGRLAYLP-----VGRYSKTPASPV--- 227
DB 198 KWRKSLGHHRTVMGFIRSCNLRSTYKGRILYRPRKGFHRSSNVFSYEXTTQQRIDDS 257
QY 228 -VVOQGPDVAHLVPLEEPVPSHW-----TVVPDEDEFLVIALLSHSL 268
DB 258 KVKTNQSVSDEEETMETKPFQNTLTPDSDETLAVGSSDLEETVVIDENFVNITAVTLSHI 317
QY 269 GSEMFAPRPMGRCAAGVHILFY--RAGVSRMLRLPLAMKGRHMECPYLYVVPYVA 326
DB 318 AADGPFAPASAKLEDNRHLSYIIMKDKDIGTRVNIATKYLAIETHL--DLPFVAKHVEVSS 375
QY 327 FRLEPKDGKGVFAVDGELMVEAVQ 351
DB 376 MKLEVISGSHVLDGEVVDKRTIE 400

```

Search completed: May 9, 2003, 16:54:15
Job time : 39 secs

Db 616 Q 616

RESULT 13

US-09-817-676A-14 Application US/09817676A

Sequence 14 Patent No. US20020042101A1

GENERAL INFORMATION: Spiegell, Sarah

APPLICANT: Kohama, Takatani

TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,

FILE REFERENCE: 00170/HG

CURRENT FILING DATE: 2001-03-26

PRIOR APPLICATION NUMBER: US 60/194,318

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 14

LENGTH: 618

TYPE: PRT

ORGANISM: Homo sapiens

US-09-817-676A-14

Query Match 44.0%; Score 886.5; DB 10; Length 618;

Best Local Similarity 39.8%; Pred. No. 4,3e-78;

Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VPRPCRYLVLLNFRGCKKALQLFPSHVOPPLAELISFTLMTERNHARELVSEEL 69
 Db 140 LPRPRRLLLVNFPGCRGLAWQCKNHVLPMSISAGLSFNLQTERQHARELVQGLSL 199
 QY 70 GRWDALVYMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEV 129
 Db 200 SEMDGIIVYSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEV 259
 QY 130 TNEDELITNCTLLCRLLSPMNLISLTASGLRFSVSLAMGFIADVDESERYRLGE 189
 Db 260 LGLDILLNCSLLCRGGCHPLDLSTVTLASGSRCSFSLVAMGFVSDVIOSEFRALDS 319
 QY 190 MRFITGFLRLAALRYRGLAYLPVG-RVGSKTPA----- 224
 Db 320 AEFITGLVGLATLHTYRGLSYLPATVBPASPTPAHSLPRAKSELTLTPDPAPMANSP 379
 QY 225 ----- 224
 Db 380 LHRVSDLPRLPQALASGSPERPLRLSLNGGRELADMGAGADAPLSPDLLSSPP 439
 QY 225 -----SPVYVQO-----GPVDALVLEEPVSHWTVVP 253
 Db 440 GSPKALHSPVSEGAIPVPSGLPLPTPDARVAGSTGCPDHLPLPGLPPLPDDN-VTL 498
 QY 254 DEDFVLYLALHSHLSEMEFAPMRCAGVNHLYFVAGVSRAMLRLFLAMEKGRHME 313
 Db 499 EEDFVLYLALHSHLSEMEFAPMRCAGVNHLYFVAGVSRAMLRLFLAMEKGRHME 558
 QY 314 YECRYLYVVPVAFRLPKDKGVAFVNDGELMVSFAVGOVHPNYFMVSGCVPEPPSWK 373
 Db 559 LCCPOLGYAARARLRLPRGVLTVDEGEYGRLOAOMHPIGTLTG----PPGC- 613
 QY 374 PQOMPPEEP 383
 Db 614 -----PGREP 618

RESULT 14

US-09-970-516-4 Application US/09970516

Sequence 4 Patent No. US2002009929A1

GENERAL INFORMATION: PatentIn Ver. 2.0

APPLICANT: NO. US2002009929A1arcls AG

TITLE OF INVENTION: Induction of blood vessel formation through administration of
 FILE REFERENCE: 4-31617
 CURRENT FILING DATE: 2001-10-04
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Version 3.1
 SEQ ID NO 4
 LENGTH: 618
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-970-516-4

Query Match 44.0%; Score 886.5; DB 10; Length 618;

Best Local Similarity 39.8%; Pred. No. 4,3e-78;

Matches 195; Conservative 61; Mismatches 107; Indels 127; Gaps 6;

QY 10 VPRPCRYLVLLNFRGCKKALQLFPSHVOPPLAELISFTLMTERNHARELVSEEL 69
 Db 140 LPRPRRLLLVNFPGCRGLAWQCKNHVLPMSISAGLSFNLQTERQHARELVQGLSL 199
 QY 70 GRWDALVYMSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEV 129
 Db 200 SEMDGIIVYSGDGLMHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGYEV 259
 QY 130 TNEDELITNCTLLCRLLSPMNLISLTASGLRFSVSLAMGFIADVDESERYRLGE 189
 Db 260 LGLDILLNCSLLCRGGCHPLDLSTVTLASGSRCSFSLVAMGFVSDVIOSEFRALDS 319
 QY 190 MRFITGFLRLAALRYRGLAYLPVG-RVGSKTPA----- 224
 Db 320 AEFITGLVGLATLHTYRGLSYLPATVBPASPTPAHSLPRAKSELTLTPDPAPMANSP 379
 QY 225 ----- 224
 Db 380 LHRVSDLPRLPQALASGSPERPLRLSLNGGRELADMGAGADAPLSPDLLSSPP 439
 QY 225 -----SPVYVQO-----GPVDALVLEEPVSHWTVVP 253
 Db 440 GSPKALHSPVSEGAIPVPSGLPLPTPDARVAGSTGCPDHLPLPGLPPLPDDN-VTL 498
 QY 254 DEDFVLYLALHSHLSEMEFAPMRCAGVNHLYFVAGVSRAMLRLFLAMEKGRHME 313
 Db 499 EEDFVLYLALHSHLSEMEFAPMRCAGVNHLYFVAGVSRAMLRLFLAMEKGRHME 558
 QY 314 YECRYLYVVPVAFRLPKDKGVAFVNDGELMVSFAVGOVHPNYFMVSGCVPEPPSWK 373
 Db 559 LCCPOLGYAARARLRLPRGVLTVDEGEYGRLOAOMHPIGTLTG----PPGC- 613
 QY 374 PQOMPPEEP 383
 Db 614 -----PGREP 618

RESULT 15

US-09-784-810A-15 Application US/09784810A

Sequence 15 Patent No. US20020082203A1

GENERAL INFORMATION: RASTELLI, LUCA

TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

FILE REFERENCE: 10716-08

CURRENT FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: 60/182,360

PRIOR FILING DATE: 2000-03-22

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 415

Best Local Similarity 81.4%; Pred. No. 7.4e-151;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

```

OY 7 PGCVLPKPCRYLVLLNPGCGKALQFPRSHVOPLLAEAEISFTLMTERRNNAARELVRS 66
Db 6 PRLLEPRPCRYLVLLNPGCGKALQFOSRVOPLFEAEITFKILTERKNHARELVCA 65
OY 67 EELGRWDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSIPAGSGNALAASVNHAGY 126
Db 66 EELGHMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSIPAGSGNALAASVNHAGY 125
OY 127 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSIACGFIADVDLESKYRR 186
Db 126 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSIACGFIADVDLESKYRR 185
OY 187 LCEMRFTLCTFLRLAALFTYRGRLAYLPVGRVSKTPASPVVVOGCPVDAHLVLEERP 246
Db 186 LCEIRFTVCTFLRLASLRIYQGLAYLPVGTASKRPAS-TLVQKGPVDTLHVPLEERP 244
OY 247 SHMTVPDEDFVLVLLSHLSGSEMFAPMGRCAAGVMHLFTYRAGVSRAALLRLFLAM 306
Db 245 SHMTVPDEDFVLVLLSHLSGSEMFAPMGRCEAGVMHLFTYRAGVSRAALLRLFLAM 304
OY 307 EKGRIMEYECPLYVYVVAFAFLRPEKDGKVFVADGELMVAVOGVHNPYFMVSGCV 366
Db 305 OKKHMELDCPLYVHVPVAFRLRPSRGVFSVDGELMVAVOGVHNPYFMVSGCV 364
OY 367 EPPPSWKPOQMPPEEP 383
Db 365 DAPSGRDSRRGPPPEEP 381

```

RESULT 9

```

US-09-796-487-4
; Sequence 4, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(381)
; OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mspk1a in Figure 3, corre
; OTHER INFORMATION: Spouding to amino acid residue 124 to 504 of SPK1a of Genbank se
; OTHER INFORMATION: Sequence Accession Number AF068748. SEQ ID NO 4 is equivalent to SE
; OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
; AUTHORS: Kohama et al.,
; TITLE: Molecular cloning and functional characterization of murine sphingosine
; TITLE: Kinase
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 273
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697

```

DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

```

Query Match 80.7%; Score 1627.5; DB 10; Length 381;
Best Local Similarity 81.4%; Pred. No. 7.4e-151;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

OY 7 PGCVLPKPCRYLVLLNPGCGKALQFPRSHVOPLLAEAEISFTLMTERRNNAARELVRS 66
Db 6 PRLLEPRPCRYLVLLNPGCGKALQFOSRVOPLFEAEITFKILTERKNHARELVCA 65
OY 67 EELGRWDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSIPAGSGNALAASVNHAGY 126
Db 66 EELGHMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSIPAGSGNALAASVNHAGY 125
OY 127 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSIACGFIADVDLESKYRR 186
Db 126 EGYTNEDLLTNCITLLCRRLSPMNLISHTASGLRFSVLSIACGFIADVDLESKYRR 185
OY 187 LCEMRFTLCTFLRLAALFTYRGRLAYLPVGRVSKTPASPVVVOGCPVDAHLVLEERP 246
Db 186 LCEIRFTVCTFLRLASLRIYQGLAYLPVGTASKRPAS-TLVQKGPVDTLHVPLEERP 244
OY 247 SHMTVPDEDFVLVLLSHLSGSEMFAPMGRCAAGVMHLFTYRAGVSRAALLRLFLAM 306
Db 245 SHMTVPDEDFVLVLLSHLSGSEMFAPMGRCEAGVMHLFTYRAGVSRAALLRLFLAM 304
OY 307 EKGRIMEYECPLYVYVVAFAFLRPEKDGKVFVADGELMVAVOGVHNPYFMVSGCV 366
Db 305 OKKHMELDCPLYVHVPVAFRLRPSRGVFSVDGELMVAVOGVHNPYFMVSGCV 364
OY 367 EPPPSWKPOQMPPEEP 383
Db 365 DAPSGRDSRRGPPPEEP 381

```

RESULT 10

```

US-09-796-487-5
; Sequence 5, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(373)
; OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in Fig. 2,
; OTHER INFORMATION: corresponding to amino acid residue 131 to 504 of SPK1a of Genb
; OTHER INFORMATION: sequence Accession Number AAC61697.
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (132)..(504)
; US-09-796-487-5

```

Query Match 79.8%; Score 1609.5; DB 10; Length 373;
Best Local Similarity 81.3%; Pred. No. 4.1e-149;
Matches 304; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

Qy	Db	Qy	Db
301	301	361	361
R.LFLAMEKGRMECECYLYVYVPAVARLREPOCKGVFAVDELMSAEAVOGQVHPYFM	R.LFLAMEKGRMECECYLYVYVPAVARLREPOCKGVFAVDELMSAEAVOGQVHPYFM	RVSGCVPEPPSPMKFOOMPPEEPPL	RVSGCVPEPPSPMKFOOMPPEEPPL
301	301	361	361
R.LFLAMEKGRMECECYLYVYVPAVARLREPOCKGVFAVDELMSAEAVOGQVHPYFM	R.LFLAMEKGRMECECYLYVYVPAVARLREPOCKGVFAVDELMSAEAVOGQVHPYFM	RVSGCVPEPPSPMKFOOMPPEEPPL	RVSGCVPEPPSPMKFOOMPPEEPPL

RESULT 2
US-09-796-487-3

Query Match	99.6%	Score 2008	DB 10	Length	384			
Best Local Similarity	99.2%	Pred. No.	4.6e-188					
Matches 381	Conservative	2	Mismatches	1	Indels	0	Gaps	0

QY	241	LEBPVSHMTVPPEDEVLVALLHSLTSGMFAPMGRCAQVMHLEFYVRAGYSRAML	3000
QY	241	LEBPVSHMTVPPEDEVLVALLHSLTSGMFAPMGRCAQVMHLEFYVRAGYSRAML	3000
Db	241	LEBPVSHMTVPPEDEVLVALLHSLTSGMFAPMGRCAQVMHLEFYVRAGYSRAML	3000
QY	301	RLFLAMEGRMEVECYLYVVPVAFRLBPKDGKGVAVDGLMTVSEAVOGVHPRYFW	3600
Db	301	RLFLAMEGRMEVECYLYVVPVAFRLBPKDGKGVAVDGLMTVSEAVOGVHPRYFW	3600
QY	361	MVSGCVERPPSMKQOMPPEEPL	384
Db	361	MVSGCVERPPSMKQOMPPEEPL	384

```

RESULT 3
US-09-784-810A-2
; Sequence 2, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-2

```

Query Match	97.38;	Score 1962;	DB 10;	Length 384;
Best Local Similarity	97.78;	Pred. No. 1.4e-183;		
Matches 375;	Conservative	2;	Mismatches 7;	Indels 0;
			Gaps	0

RESULT 4
US-09-784-810A-4
; Sequence 4, Application US/09784810A
; Patent No. US20020082203A1

Mon May 16 10:50:39 2003

us-09-937-060a-5.rapb

Page 1

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:49:40 ; Search time 23 Seconds
(without alignments)
1536.429 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPAGCGVLPKRCVLYL.....CPEPPSKWKROOMPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2016	100.0	384	10	US-09-970-516-2
2	2008	99.6	384	10	US-09-796-487-3
3	1962	97.3	384	10	US-09-784-810A-2
4	1639.5	81.3	382	10	US-09-784-810A-4
5	1632.5	81.0	382	10	US-09-970-516-6
6	1630.5	80.9	388	10	US-09-817-676A-15
7	1630.5	80.9	388	10	US-09-796-487-2
8	1627.5	80.7	381	10	US-09-796-487-1
9	1627.5	80.7	381	10	US-09-796-487-5
10	1609.5	79.6	373	10	US-09-796-487-4
11	938.5	44.0	618	10	US-09-817-676A-12
12	893	44.3	617	10	US-09-817-676A-14
13	886.5	44.0	618	10	US-09-970-516-1
14	886.5	44.0	618	10	US-09-784-810A-15
15	374.5	18.6	424	10	US-09-796-487-8
16	374	18.6	424	10	US-09-784-810A-12
17	356.5	17.7	453	10	US-09-784-810A-11
18	356.5	17.7	453	10	US-09-784-810A-14
19	355	17.6	380	10	US-09-784-810A-14

20	345	17.1	403	10	US-09-796-487-7	Sequence 7, Appl
21	345	17.1	436	10	US-09-784-810A-13	Sequence 13, Appl
22	251	12.5	471	10	US-09-784-810A-6	Sequence 6, Appl
23	227.5	11.3	326	10	US-09-784-810A-11	Sequence 11, Appl
24	160	7.9	365	9	US-09-738-626-6336	Sequence 29, Appl
25	159	7.9	365	9	US-09-784-810A-29	Sequence 29, Appl
26	155.5	7.7	421	10	US-09-948-820-50	Sequence 50, Appl
27	147	7.3	240	10	US-09-796-487-10	Sequence 10, Appl
28	124	6.2	296	10	US-09-815-242-4883	Sequence 4883, Ap
29	124	6.2	362	10	US-09-815-242-10793	Sequence 10793, A
30	111	5.5	294	10	US-09-815-242-13334	Sequence 13334, A
31	106	5.3	315	10	US-09-815-242-12882	Sequence 12882, A
32	100	5.0	315	10	US-09-815-242-5505	Sequence 5505, Ap
33	100	5.0	315	10	US-09-815-242-12494	Sequence 12494, A
34	99.5	4.9	1085	10	US-09-771-161A-239	Sequence 239, App
35	99	4.9	3782	9	US-09-860-846-4	Sequence 4, Appl
36	99	4.9	3782	9	US-09-968-848-4	Sequence 4, Appl
37	99	4.9	3782	10	US-09-861-289-4	Sequence 4, Appl
38	92	4.6	782	9	US-09-908-193-47	Sequence 47, Appl
39	92	4.6	1429	10	US-09-996-617-2	Sequence 2, Appl
40	92	4.6	1429	10	US-09-931-071-2	Sequence 2, Appl
41	92	4.6	1473	9	US-09-388-221-2	Sequence 2, Appl
42	91	4.5	1399	9	US-09-388-221-4	Sequence 4, Appl
43	91	4.5	1443	9	US-09-388-221-6	Sequence 6, Appl
44	89.5	4.4	341	9	US-09-738-626-3536	Sequence 3536, Ap
45	89.5	4.4	974	9	US-10-101-464A-921	Sequence 921, App

ALIGNMENTS

RESULT 1
US-09-970-516-2
Sequence 2, Application US/0970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT FILING DATE: 2001-10-04
CURRENT APPLICATION NUMBER: US/09/970, 516
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-2,
Query Match 100.0%: Score 2016; DB 10: Length 384;
Best Local Similarity 100.0%: Pred No. 7.5e-189;
Matches 384; Conservative 0; Mismatches 0; Indels 0;

QY	1	MDPAGCGVLPKRCVLYLNLPRGKKAQLQFRSHVOPVLAEISTLMLTRRNA	60
DB	1	MDPAGCGVLPKRCVLYLNLPRGKKAQLQFRSHVOPVLAEISTLMLTRRNA	60
QY	61	RELVRSEELGRMDALVYNSGDIHMEVYNGMERPDWTAIQPLCSLPAGSGNALASL	120
DB	61	RELVRSEELGRMDALVYNSGDIHMEVYNGMERPDWTAIQPLCSLPAGSGNALASL	120
QY	121	NHYAGYQDVNEDILNCTLLCRLLSPMNLISLTAASGLRFLSVLSLAWGFIDVLE	180
DB	121	NHYAGYQDVNEDILNCTLLCRLLSPMNLISLTAASGLRFLSVLSLAWGFIDVLE	180
QY	181	SEKVRRLGEMKFTLGFRLALRTYRGRVAVGSKTPASPVVVOQCPVAHLVP	240
DB	181	SEKVRRLGEMKFTLGFRLALRTYRGRVAVGSKTPASPVVVOQCPVAHLVP	240
QY	241	LEEPVPSHWTVVDEDFVLYLALSHSGSEMPAAPMGRCACAAVYMLFYRAGVSRAML	300
DB	241	LEEPVPSHWTVVDEDFVLYLALSHSGSEMPAAPMGRCACAAVYMLFYRAGVSRAML	300

Mon May 1~~2~~ 10:50:39 2003

us-09-937-060a-5.rai

Page 7

[illegible]

Search completed: May 9, 2003, 16:50:25
Job time : 28 secs

APPLICANT: Cosman, David J.
APPLICANT: Anderson, Dirk M.
APPLICANT: Borges, Luis
TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-Like Receptors (LIR)
FILE REFERENCE: 2624-A
CURRENT APPLICATION NUMBER: US/09/310,463A
CURRENT FILING DATE: 1999-05-12
EARLIER APPLICATION NUMBER: 08/842,248
EARLIER FILING DATE: 1997-04-24
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 590
TYPE: PRT
ORGANISM: human
US-09-310-463-22

Query Match 4.3%; Score 86; DB 4; Length 590;
Best Local Similarity 24.8%; Pred. No. 0.86;
Matches 62; Conservative 21; Mismatches 79; Indels 88; Gaps 15;

QY 162 EKYRLGEMRFL-----GTFRLAALRTYGRGLAYLPVGVGSKTPASP-VVVOQGV 234
DB 372 OSYRHOAEFMSPTVSAOGTCYSAIRSY-----PYL-----LSSPSYPOELVVGSPS 421
QY 235 -DAHLVPL-EEVPSSHMTVPDEDFVLVAL-----LHSHLGSEMFAPMGCAAGVMHL 287
DB 422 GPPSLSPGSGTTPG-----PEDQPLTPGLDPQSLGRHLG----- 458
QY 288 FVVRAGVSRAMLRLFL--AMEKGRHM-EYECPLYLVVPPVVAFLRLEPKD----- 333
DB 459 --VVTGVSVAFLVLLFLFLRLRHRHOSKHTSAHFYRPAGAAGPEPKDQLOKRASPV 516
QY 334 -----GKGVFAVDGELMYSEAVQ-----GQVPRNTFMWVSGCVERPPSPK 373
DB 517 ADIQEELNAAVKDTQPKDGVEMDARAASAPQDVTYAQLHS--LTLREATTEPPS-- 572
QY 374 PQMPPEEP 383
DB 573 -QERREPAEP 581

RESULT 14
US-08-842-248A-22
Sequence 22, Application US/08842248A
Patent No. 6448035
GENERAL INFORMATION:
APPLICANT: Cosman, David J.
TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-Like Receptors (LIR)
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM/PC Compatible
OPERATING SYSTEM: Microsoft Word 7.0
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,248A
FILING DATE: April 24, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C.
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2624
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 590 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-842-248A-22

Query Match 4.3%; Score 86; DB 4; Length 590;
Best Local Similarity 24.8%; Pred. No. 0.86;
Matches 62; Conservative 21; Mismatches 79; Indels 88; Gaps 15;

QY 162 EKYRLGEMRFL-----GTFRLAALRTYGRGLAYLPVGVGSKTPASP-VVVOQGV 234
DB 372 OSYRHOAEFMSPTVSAOGTCYSAIRSY-----PYL-----LSSPSYPOELVVGSPS 421
QY 235 -DAHLVPL-EEVPSSHMTVPDEDFVLVAL-----LHSHLGSEMFAPMGCAAGVMHL 287
DB 422 GPPSLSPGSGTTPG-----PEDQPLTPGLDPQSLGRHLG----- 458
QY 288 FVVRAGVSRAMLRLFL--AMEKGRHM-EYECPLYLVVPPVVAFLRLEPKD----- 333
DB 459 --VVTGVSVAFLVLLFLFLRLRHRHOSKHTSAHFYRPAGAAGPEPKDQLOKRASPV 516
QY 334 -----GKGVFAVDGELMYSEAVQ-----GQVPRNTFMWVSGCVERPPSPK 373
DB 517 ADIQEELNAAVKDTQPKDGVEMDARAASAPQDVTYAQLHS--LTLREATTEPPS-- 572
QY 374 PQMPPEEP 383
DB 573 -QERREPAEP 581

RESULT 15
US-09-144-085-2
Sequence 2, Application US/09144085
Patent No. 6280999
GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziermann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
FILE REFERENCE: 30062-20020, 20
CURRENT APPLICATION NUMBER: US/09/144,085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010,809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 6095
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match 4.3%; Score 86; DB 4; Length 6095;
Best Local Similarity 23.4%; Pred. No. 33;
Matches 64; Conservative 29; Mismatches 82; Indels 98; Gaps 12;

QY 1 MDPAGPGRVLPKPCVLYLVNDRGSKKALDIFRSHVPLAEAFISTMLTERRN-- 58
DB 1427 LDPAG-----TVLV-----TGSGTGLGQAIH-----LVRAHGVHVLVTSRGL 1468
QY 59 --HARELYRS-FELGRWDALVWVSGDGLMHVYVNGLMERPDMETALOKLSPASGNA 115
DB 1469 APCARELYVSLBELGA-ETVVAACDVSKREEVARVLAGID--AARPL----- 1513

Mon May 1 10:50:39 2003

us-09-937-060a-5.ra1

Page 5

```
Db 11184 GAF-DLDE-----GRSAVLAMVWMOASVJRCDSPEEYVDPDLRFEGDSLTAVDFR 11237
QY 208 GHLAVLPVGRVSKTPASPVYVOOGPVDAHNP---LEPVSHTVTPDEDFVLVALLH 265
Db 11238 NKVNL-----TGDLPP--PVYVQHPPTVALAERISDLAERNAVAPESD-----H 11283
QY 266 SHLSEMFMAAMGR-----CAQVMHLFVYRAGVSRAMLLRLFLAMKGRHMY 314
Db 11284 EGAERKAAPAGARSGADTGAQ-----AGMERLFRQ--AYEDDYGEEF 11326
```

```
RESULT 11
US-09-320-878-18
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06 OF 08/046,247
; EARLIER APPLICATION NUMBER: CIP OF 08/119,139
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1999-02-08
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRP
; ORGANISM: Streptomyces venezuelae
US-09-320-878-18
```

```
Query Match 4.3%: Score 86; DB 3: Length 416;
Best Local Similarity 22.2%; Pred. No. 0.5;
Matches 94; Conservative 51; Mismatches 129; Indels 150; Gaps 23;
```

```
QY 19 VLNPRGKKAQLQFRSHVQPL-AEAISFTLMLTERRNHR-----EL 63
Db 64 VLADPRSKD-----WNSTPTLFEALNHNMLSDPFRHTRKLKLVAREFTWRVEL 118
QY 64 VRSEELGRMDALV---VMSGDG---LMHE-----VYNGLMERPDP-----METAI 101
Db 119 LRPRVOETIVDGLDAMLAADGADLMESLAMPPLPTIVISLGLVPEPRAAFRWTDAT 178
QY 102 QKPLCSLPAGSGNALASLNHYAGY-----EYVNTED 133
Db 179 VFP--DDPAQAQOTMA---EMSGYLSRLIDSKRGODGEDLSALVRTSDEDSRLTSEE 232
QY 134 LITNCTLLCRRLSPMNLISLHTASGLRFSVLAMGFADVDLSEKRYRRLGEMRPT 193
Db 233 LLGMAHILVAGHETVNI-----ANGM-----YALLSHPD-----264
QY 194 LGTFLRLAALRTYGRALYLPVGRVSKTPASPVYVOOGPVDA--HLVPLEPVPSHMTV 251
Db 265 -----QLAALRADMTLLD---GAV-----EEMLRVGEPEESATYREFV--EPVLDLGTV 308
QY 252 VDDPEFVLVALLSHLSEMFMAAMGRCAAGVNHLPYRAGVSRAMLLRLFLAMKGRH 311
Db 309 IPAGDTVLVY-LADAHRTPERFPDP-----HRFDJFRDTAG-----HLAFGHIGH 352
```

```
QY 312 MEYECPLVYVPPVAFR--LEPKDCKGVFAVDGELMAYSEAVOGVHPNFMVSGCVERP 369
Db 353 FCICAPLARLEAIAVALLERCPDLALDVSFGLV-----WTFNP--MIRKLALP 402
QY 370 PSKR 373
Db 403 IRWR 406
```

```
RESULT 12
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRP
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39
```

```
Query Match 4.3%: Score 86; DB 4: Length 416;
Best Local Similarity 22.2%; Pred. No. 0.5;
Matches 94; Conservative 51; Mismatches 129; Indels 150; Gaps 23;
```

```
QY 19 VLNPRGKKAQLQFRSHVQPL-AEAISFTLMLTERRNHR-----EL 63
Db 64 VLADPRSKD-----WNSTPTLFEALNHNMLSDPFRHTRKLKLVAREFTWRVEL 118
QY 64 VRSEELGRMDALV---VMSGDG---LMHE-----VYNGLMERPDP-----METAI 101
Db 119 LRPRVOETIVDGLDAMLAADGADLMESLAMPPLPTIVISLGLVPEPRAAFRWTDAT 178
QY 102 QKPLCSLPAGSGNALASLNHYAGY-----EYVNTED 133
Db 179 VFP--DDPAQAQOTMA---EMSGYLSRLIDSKRGODGEDLSALVRTSDEDSRLTSEE 232
QY 134 LITNCTLLCRRLSPMNLISLHTASGLRFSVLAMGFADVDLSEKRYRRLGEMRPT 193
Db 233 LLGMAHILVAGHETVNI-----ANGM-----YALLSHPD-----264
QY 194 LGTFLRLAALRTYGRALYLPVGRVSKTPASPVYVOOGPVDA--HLVPLEPVPSHMTV 251
Db 265 -----QLAALRADMTLLD---GAV-----EEMLRVGEPEESATYREFV--EPVLDLGTV 308
QY 252 VDDPEFVLVALLSHLSEMFMAAMGRCAAGVNHLPYRAGVSRAMLLRLFLAMKGRH 311
Db 309 IPAGDTVLVY-LADAHRTPERFPDP-----HRFDJFRDTAG-----HLAFGHIGH 352
QY 312 MEYECPLVYVPPVAFR--LEPKDCKGVFAVDGELMAYSEAVOGVHPNFMVSGCVERP 369
Db 353 FCICAPLARLEAIAVALLERCPDLALDVSFGLV-----WTFNP--MIRKLALP 402
QY 370 PSKR 373
Db 403 IRWR 406
```

```
RESULT 13
US-09-310-463-22
; Sequence 22, Application US/09310463A
; Patent No. 6384203
; GENERAL INFORMATION:
```

OY 186 RLGEKRTLTGFLRLALRYRGLAVLPVG--RVGSKTPASPVVQ-----QGPV 234
DB 853 ALGTGPHLDCSLRWLSWVAKG---YKEPGIARCSSEPMADRLLTPTPHRQCKGPV 909
OY 235 DAHLVP-----LEPPVPSHMTVVPDEDPVVLVLLHSHLGSEMPAARGRA-----AG 283
DB 910 DINIVAKCNACLSPPCKNNGTCTODPVELYRCACPYSYKQDC--TVPLNTICQNPCHG 968

OY 284 VMHL 287
DB 969 TCHL 972

RESULT 8

US-09-105-537-37
Sequence 37, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 37
LENGTH: 1346
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-37

Query Match 4.3%; Score 86.5; DB 4; Length 1346;
Best Local Similarity 22.1%; Pred. No. 2.7;
Matches 65; Conservative 17; Mismatches 81; Indels 131; Gaps 14;

OY 81 DGLMEVYVNGLMERPDMETAIQKPLCSL-----PAGSG 113
DB 854 DGGHRLTTSIAEA--WANGIALDWASLLPATGALSPAVPDLPYAFQHRSYWISPAAG 911
OY 114 NALAASLNHYAGYEQVTNEDLLTNTCLLRCRLSPMNLISHTASGLRFSVLSIANGF 173
DB 912 EAPA-----HTASGREVAETGLANGP 933
OY 174 IADVLESEKRYRIGEMRFTLTGFLRLA-----Lr-----TYR 207
DB 934 GAE-DLDBE---GRSAVILAMVMOAASVLRCDSPPEVPVDRPLREIGEDSLTAVDFR 987
OY 208 GLAYLPVGRVSKTPASPVVVOGSPVDAHLVP--LEPPVPSHMTVVPDEDPVVLVLLH 265
DB 988 NRVNLL-----TGLQLP--PTYVFQHPPTVALAERISDELAEKNNMVAEPSP-----H 1033
OY 266 SHLGSEMPAARGRA-----CAAGVMHLFYRVAGVSRAMLRLFLAMEGRHMEY 314
DB 1034 EQAEERKAAPAGANSAGADTGAG-----AGMFRALFRQ--AVEDDRYGEF 1076

RESULT 9

US-09-245-041-15
Sequence 15, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
APPLICANT: Nagle, D.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245.041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21

EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 15
LENGTH: 2787
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-041-15

Query Match 4.3%; Score 86.5; DB 4; Length 2787;
Best Local Similarity 22.4%; Pred. No. 8.5;
Matches 68; Conservative 34; Mismatches 101; Indels 101; Gaps 14;

OY 97 WETAIQKPLCSLPAGSGNALASLNHYAG-----YEQVTNEDLLTNTCLL 142
DB 2484 WEAFYLPVOLLFLFNGYVMOGLHIFGRMWCHEHLSHFHTSIHSTSI---CTITA 2540
OY 143 CRRLSPMNLISL-----HTASGLRFSVLSIANGFTADVLESEKRYRIGEMRFTLTG 196
DB 2541 NARVYTEKILKONPSHPVP--IPGDLSTLVW----- 2572
OY 197 FLRLALTYTGR-----LAYLPVGRVSKTPASPVVVOGSPVDAHLVPLEE 243
DB 2573 FLEALILRTPQGRORPPLGRACPAVLLKCLRPNL--APAIAYVSRGP--PRDIAPESN 2628
OY 244 PVPSSHMTVVPDE-----DFVLVLLHSHLGSEMPAARGCAAGVMHLF-- 288
DB 2629 TLPTRCFEMERKIGTSSFKSGNSGIDPOLNICTKRSK--HSLFRNSSGAVFYL-LFTI 2665
OY 289 --YVRAGVSRAML-----LRLFLAMEGRHMEYCPYLV-YVPAVRLEPKD 333
DB 2686 SEVWIOAEFVNMNVLASITLIMISKYCKLFHCGDCGCTYIDCNFLFVFLVNNSPFOEN 2745
OY 334 GKGV 337
DB 2746 GRSV 2749

RESULT 10

US-09-105-537-6
Sequence 6, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-6

Query Match 4.3%; Score 86.5; DB 4; Length 11877;
Best Local Similarity 22.1%; Pred. No. 82;
Matches 65; Conservative 17; Mismatches 81; Indels 131; Gaps 14;

OY 81 DGLMEVYVNGLMERPDMETAIQKPLCSL-----PAGSG 113
DB 11104 DGGHRLTTSIAEA--WANGIALDWASLLPATGALSPAVPDLPYAFQHRSYWISPAAG 11161
OY 114 NALAASLNHYAGYEQVTNEDLLTNTCLLRCRLSPMNLISHTASGLRFSVLSIANGF 173
DB 11162 EAPA-----HTASGREVAETGLANGP 11183
OY 174 IADVLESEKRYRIGEMRFTLTGFLRLA-----Lr-----TYR 207

Mon May 10:50:39 2003

us-09-937-060a-5.rai

Page 3

Db 564 TTAELLVORFASIDELAAATIDELAALEGVGPITAEISIANMFVEDNRRLIEELKEIGVN 623
QY 241 ---LEEPSPSHMTVPDEVLVIALH--SHLSEMFAPMGCAAGV 284
Db 624 TQRLPEAPAEAPVKGKTFVLGALPHILTRKEABELIKRAGGVASV 672

RESULT 5

US-08-936-135-4
Sequence 4, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-4

Query Match 4.5%; Score 91; DB 3; Length 2584;
Best Local Similarity 22.5%; Pred. No. 2.3;
Matches 67; Conservative 44; Mismatches 123; Indels 64; Gaps 12;

QY 41 LLAELISFTMLTERNNHRE--LYRSEELGRDALVMS--GDGLMEVYNGLMERP 95
Db 2049 LYSERHIVAGLITTHLEKRVALLYSLEHISTYRGIANLISPGIGITVRASPGI----- 2104
QY 96 DWETAIQPLCSLPAGSGNALASLNHYAGT-----QVTNEDLITNCTLLCRRLSP 149
Db 2105 NLEVALTPMPTVA-----LVALLGTHISGLNGLIASPHISITRPLYSGLDLVARGVALL 2159
QY 150 NMLLSLHATASGLRFSVLSLAWGFIADVLESEKYRIQ-----EMFTTIGTF-LIRLAI. 203
Db 2160 LKHISYSERLLEYSL-----ETVRGLNVALILPEHREGIGLGLTLEGL 2204
QY 204 RTYRGRLA-----LTPGVGSKTPASPVVVOGQPYDA-----HLVLEEPVPSHM 249
Db 2205 YLISGLYASNEGLGYGLYIDEALVALASPASPILESERITLSEASNNHISILEPGLNG 2264
QY 250 TVVPEDEVVLALLSHLSGSEMFAPMGCAAGVNHLYFVRACVSRAMLRLFLAME 307
Db 2265 LASFYS-ALALYSPRTRHASP-----LysLYSASNTHRGLLELYSTILE 2313

RESULT 6
US-09-134-001C-5553
Sequence 5553, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5553
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5553

Query Match 4.3%; Score 87.5; DB 4; Length 319;
Best Local Similarity 23.7%; Pred. No. 0.22;
Matches 28; Conservative 26; Mismatches 47; Indels 17; Gaps 4;

QY 22 NPRGKGKRALDTRSHVOPLAE--AEISFTMLTERNNHARELYRSEELGRDALVMS 79
Db 35 NGIGEVAKSLSSMCKHLSLQSENKGLIKYCKSIKENTSSDV-----DVLFIIG 85
QY 80 GDGLMEVYNGLMERPDMETAIQPLCSLPAGSGNALASLNHYAGTGYTMDLLTN 137
Db 86 GDGTNLELVNGVM-----QYQLNPLGVIPGTFNDEFTKTLQHPN-RTASEQLTIS 137

RESULT 7
US-09-182-024A-2
Sequence 2, Application US/09182024A
Patent No. 6342370
GENERAL INFORMATION:
APPLICANT: Connolly, Timothy
TITLE OF INVENTION: Human Silt Polypeptide and Polynucleotides Encoding
FILE REFERENCE: 640100-271
CURRENT APPLICATION NUMBER: US/09/182,024A
CURRENT FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/063,946
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/096,420
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1523
TYPE: PRT
ORGANISM: Homo sapiens
US-09-182-024A-2

Query Match 4.3%; Score 87; DB 4; Length 1523;
Best Local Similarity 24.2%; Pred. No. 2.9;
Matches 59; Conservative 26; Mismatches 101; Indels 58; Gaps 10;

QY 98 ETATQ--KPLCSLPAG-----SGNALASLNHYAGTGYTMDLLTNCTLLCR 144
Db 733 ETVVRCSNKGRLALPGMKDVTETLYLGNHITVAPRELALRHITLIDLSNNSISMLTN 792
QY 145 RLSPMLLS-----LHPASGLRFSVLSLAWGFIADV-----DLESEKYR 185
Db 793 YTFSMHISTLILSYNRRLRCIPVHAFNGLSLAVLITHGNDISSVPBSGFNDLISLH 852

QY 353 QVBNPY 358
Db 285 KLPANF 290

RESULT 2

US-09-412-545-2
Sequence 2, Application US/09412545
Patent No. 6255095
GENERAL INFORMATION:
APPLICANT: Prescott, Stephen M.
APPLICANT: Ding, Li
TITLE OF INVENTION: HUMAN DIACYLGLYCEROL KINASE IOTA
FILE REFERENCE: 1321.2.25
CURRENT APPLICATION NUMBER: US/09/412,545
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1065
TYPE: PRT
ORGANISM: Homo sapiens
US-09-412-545-2

Query Match 4.9%: Score 99.5; DB 4; Length 1065;

Best Local Similarity 24.6%; Pred. No. 0.064; Matches 55; Conservative 33; Mismatches 65; Indels 71; Gaps 12;

QY 3 PAGEPRGVLPKPCRVVLNPRGKGK---KALQLFRSHVQPLAEALISFTLMTERRNH 59
Db 368 PISPP---LAMP---LVFVNPKSGNGCTKVLQKFMVWLN----- 403
QY 60 ARELYRSEELGRNAL-----VMSGDLMEHVNVGMLERPPWETAIOKPLCSL 108
Db 404 -ROYFDSLQSEPKALDELRYKVPNLRIACGDCGTVCWILSTLDE---LOLSPQPPGYL 459
QY 109 PAGSGNALAASLNHYAGYEQVTNEDLLNCTLLCR---RLISPMNLISLHTASG----- 160
Db 460 PLGTGNDLARIKLNNGGCT---TDEPV-----SKILQVEDGTGVVOLDRRNLHVERNPDLPP 512
QY 161 -----LRLF-SVLSLWAGFIADVDESEKRYRLGEMRF 192
Db 513 EELEDGVCKLEPLVNNYFSL--GDAHVTLLEPHSREANPEKF 554

RESULT 3

US-09-105-537-4
Sequence 4, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
EARLIER FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3782
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-4

Query Match 4.9%: Score 99; DB 4; Length 3782;
Best Local Similarity 24.1%; Pred. No. 0.53; Matches 103; Conservative 48; Mismatches 169; Indels 88; Gaps 19;

QY 3 PAGEPRGVLPKPCR-VLVLNPRGCKKALQLFRSHVQPLAEALISFTLMTERRNHAR 61
Db 132 PEGORAVVEGIRHRELEGLAPDPSASAYAFELLCGFVRAVTAATAAANVLGVADRADFA 191
QY 62 ELVR-----SEELGRNAL-VMSGDLMEHVNVGMLERPPWETAIOKPLCSLPG---SG 113
Db 192 DLRLRLPSLSLAPSLRTRVRAADGALAEALTALADSDSPGALSALGVTAAYOLTG 251
QY 114 NALASLNHYAGYEQVTNEDLLNCTLLCRRLISPMNL-----LSLHTASGLRL----- 163
Db 252 NAVLALLAHPRQMBELCDRLAANAABETLRYPQOLDARVYRGFETLAGRLPACAH 311
QY 164 FSVLSLWAGFIADVDESEKX---RUGEMRFTL-----GTPLRL-----AALRTYRG 208
Db 312 VVLTAAATGRDPEVFTDPERDLARPDAAHILALHPACQPVASVLYLQLEVALRTLAG 371
QY 209 RLATL-----PVGVGSKTPASP---VVQOGPDAHLVPLEEVPVSHWYVP 253
Db 372 RFPGLRQAGDVLRRRAVVGRLSVPVSSMRVLLTSFAHHTHYGGL---VPLAMALLA 428
QY 254 DEDVLYLV---ALLSHLSGEMFAAPMGCAAGVNHLEFVR-AGVSRAMLRLFLAMEKG 309
Db 429 AGHEYRASQPALDITDTTSGSLAAPVG--TDHLIHEYRVMAEGRP----- 474
QY 310 RHMEYECPLYVVPYVAF---RLERKDGKGVFAVDGELMASEAVQGVHPYFMVWSGCV 366
Db 475 NH-----PAIARDEARPELDMDHAGIEAIL-----APYFHLANNDSVVDLV 519
QY 367 EPPSWKP 374
Db 520 DPARSWOP 527

RESULT 4

US-09-708-426-9
Sequence 9, Application US/09708426
Patent No. 6444429
GENERAL INFORMATION:
APPLICANT: HAN, YE-SUN
APPLICANT: YU, YEON-GYU
APPLICANT: LIM, JAE-HWAN
TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUT
FILE REFERENCE: 199579050
CURRENT APPLICATION NUMBER: US/09/708,426
EARLIER FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: KR99-49591
PRIOR FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin version 3.0
SEQ ID NO 9
LENGTH: 712
TYPE: PRT
ORGANISM: Rhodothermus marinus
US-09-708-426-9

Query Match 4.6%: Score 93; DB 4; Length 712;

Best Local Similarity 21.8%; Pred. No. 0.19; Matches 63; Conservative 37; Mismatches 109; Indels 80; Gaps 11;

QY 59 HARELYRSEELGRMDALVMSGDLMEHVNGIM-----ERPDWETAIOKPLCS----- 107
Db 401 HNEDYVRSRDIRIGLVVYIRAGDVIPQVVRPEVYEAFTGNERP--WRMPERCPSGCSQLVR 459
QY 108 LPAGSGNALAS-----LNHYAGYEQVTNEDLLNCTLLCR--LISPM----- 150
Db 460 LPGEADYVCVADCPACQFVRLIEHFAGRDAMIEGMSGVAROLAESGLVRLSDLYRLK 519
QY 151 --NLSLHTASGLRLFSVLSLWAGFIADVDESEKRYRLGEMRFITGFLRLAALRTYRG 208
Db 520 LEDILKLKGFATRRARNLLRA-----TEASQORLSRLRLFLG-----IRHWK 563
QY 209 RLATYLPVGRVGSKTPASPVVVOO-----GPVDAHLVP----- 240

Mon May 1 10:50:39 2003

us-09-937-060a-5.ra1

Page 1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: May 9, 2003, 16:46:18 ; Search time 15 Seconds
(without alignments)
753.227 Million cell updates/sec

Title: US-09-937-060A-5

Perfect score: 2016
Sequence: 1 MDPACGPGVLPKRCVRLV.....CVEPPSMKPOMPPEEPL 364

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2.6/ptodata/1/laa/5A_COMB.pep: *
2: /cgn2.6/ptodata/1/laa/5B_COMB.pep: *
3: /cgn2.6/ptodata/1/laa/6A_COMB.pep: *
4: /cgn2.6/ptodata/1/laa/6B_COMB.pep: *
5: /cgn2.6/ptodata/1/laa/PCVUS_COMB.pep: *
6: /cgn2.6/ptodata/1/laa/PCVUS_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	5.1	324	US-09-134-001C-4774	Sequence 4774, App1
2	99.5	4.9	1065	US-09-412-545-2	Sequence 2, App1
3	99	4.9	3782	US-09-105-537-4	Sequence 4, App1
4	93	4.6	712	US-09-708-426-9	Sequence 9, App1
5	91	4.5	2584	US-08-936-135-4	Sequence 4, App1
6	87.5	4.3	319	US-09-134-001C-5553	Sequence 5553, App1
7	87	4.3	1523	US-09-182-0248-2	Sequence 2, App1
8	86.5	4.3	1346	US-09-105-537-32	Sequence 32, App1
9	86.5	4.3	2787	US-09-245-041-15	Sequence 15, App1
10	86.5	4.3	11877	US-09-320-878-18	Sequence 18, App1
11	86	4.3	416	US-09-105-537-6	Sequence 6, App1
12	86	4.3	416	US-09-105-537-18	Sequence 18, App1
13	86	4.3	416	US-09-105-537-39	Sequence 39, App1
14	86	4.3	590	US-09-310-463-22	Sequence 22, App1
15	86	4.3	6095	US-08-842-248A-22	Sequence 22, App1
16	85.5	4.2	1580	US-08-842-085-2	Sequence 2, App1
17	85.5	4.2	1580	US-08-804-227C-11	Sequence 11, App1
18	84.5	4.2	471	US-08-176-427B-11	Sequence 11, App1
19	84.5	4.2	471	US-08-356-060A-5	Sequence 5, App1
20	84.5	4.2	471	US-08-460-900C-34	Sequence 34, App1
21	84.5	4.2	471	US-08-757-230X-8	Sequence 8, App1
22	84.5	4.2	471	US-08-674-509B-34	Sequence 34, App1
23	84.5	4.2	471	US-08-954-69B-34	Sequence 34, App1
24	84.5	4.2	471	US-08-957-87A-34	Sequence 34, App1
25	84	4.2	423	US-08-760-797A-1	Sequence 1, App1
26	84	4.2	424	US-08-932-929B-1	Sequence 1, App1
27	83.5	4.1	1346	US-09-320-878-4	Sequence 4, App1

28	83	4.1	424	US-08-760-797A-3	Sequence 3, App1
29	83	4.1	424	US-08-932-929B-3	Sequence 3, App1
30	83	4.1	435	US-08-031-538-11	Sequence 11, App1
31	83	4.1	520	PCT-US95-04801-5	Sequence 5, App1
32	82.5	4.1	471	US-09-325-256-26	Sequence 26, App1
33	82	4.1	840	US-08-500-857A-6	Sequence 6, App1
34	80.5	4.0	287	US-07-952-817-24	Sequence 24, App1
35	80.5	4.0	287	US-08-462-467B-2	Sequence 2, App1
36	80	4.0	1447	US-08-340-406-19	Sequence 19, App1
37	80	4.0	1447	US-08-656-055-19	Sequence 19, App1
38	80	4.0	1447	US-08-268-140-5	Sequence 5, App1
39	80	4.0	1447	US-09-268-140-5	Sequence 5, App1
40	80	4.0	1447	US-08-918-658-19	Sequence 19, App1
41	80	4.0	1447	PCT-US95-13233-19	Sequence 19, App1
42	79	3.9	928	US-08-841-483-4	Sequence 4, App1
43	79	3.9	928	US-09-382-911-4	Sequence 4, App1
44	79	3.9	1117	US-08-841-483-6	Sequence 6, App1
45	79	3.9	1117	US-09-382-911-6	Sequence 6, App1

ALIGNMENTS

RESULT 1
US-09-134-001C-4774
Sequence 4774, Application US/09134001C
Patient No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4774
LENGTH: 324
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4774

Query Match 5.1% Score 102: DB 4: Length 324:

Best Local Similarity 19.7% Pred No 0.0052:
Matches 72; Conservative 58; Mismatches 126; Indels 110; Gaps 18:

QY	16	RVLVLLNPGGKGLDIFRSHVOPPLAEAFSTMLTERNNHARELYRSEELG-----	70
DB	12	RAVILNPTSGK-----ELPKR-----VLPPDA-----LTKLEKAGTSAVATEKIGDAPTE	58
QY	71	-----RMDALVMSGGLMEVYVNGIMERPMEYTAICKPLCSLPAGSGNALASLNHY	123
DB	59	AERALESEYDLDLIAAGGGTLEVVNGIAEOPN-----RPKGLVTPKGTVDNDFGRALH--	111
QY	124	AGYEYVNTDLNCTLLCRLLSPNN-LLSLTASGLRFSVLSLWGFIAVDLESE	182
DB	112	-----LPSPDINGAIDVILIDHTTK-----VIGGMNN	138
QY	183	KYRRLDEMFLLGTFLLRLAALRYGRLAVLPGRVSGTSPASPVVVOGPDVAHLPLE	242
DB	139	RY-----FINLAA-----CGKLT-----QVSYETP-SKLKSIVGPF-AVYIKGF	175
QY	243	EPVPSHWTV-----VPDEDFVVLVALLSHLSGSEMFAPRGCAAGVHLFVYRA	292
DB	176	EMLPQKKAADVAVIEYDDNIPGEFALLRLGLTSMGPFKL-VPAKIDDDGYFLILYER	234
QY	293	GVSRAALLRLFLAMEKSGHMEYECFPLVYVPAFLERDKGKGFVAVDGLMYSEAVOG	352
DB	235	ANLAEGLHITMLASRGHTKH--PKVITAKASINISS-----FTDMOLNVGEGYCG	284

Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

```

OY 7 PGVLPKPCRVLYLNLNPGKGAQLOLFRSHVQPLAEAEISFTLMTERRNHARELYRS 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 PGLLPKPCRVLYLNLNPGKGAQLOLFRSHVQPLAEAEISFTLMTERRNHARELYRS 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 67 BELGRWDALVYNSGDDGLMHEVYVNGLMERPDMEATOKPICSIPAGSGNALAASINHYAGY 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 BELGHWDAVAMSQDGLMHEVYVNGLMERPDMEATOKPICSIPAGSGNALAASINHYAGY 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 127 EQVYNEDLLTNCITLLCRRLSPMNLISLHTASGLRLEFSVLSLWGFADVDLESEKTYR 186
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 EQVYNEDLLTNCITLLCRRLSPMNLISLHTASGLRLEFSVLSLWGFADVDLESEKTYR 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 187 LGEMFTLGTFLRLAALTYRGLAYLPVGRVSKTPASPVVVOGFPVDAHLVPLEEVP 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 LGEMFTLGTFLRLAALTYRGLAYLPVGRVSKTPASPVVVOGFPVDAHLVPLEEVP 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 247 SHMTVVPDEDFVLVALLSHSGSEMFAPMGRCAGVMHLFVVRAGVSRAMLLRLFLAM 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 SHMTVVPDEDFVLVALLSHSGSEMFAPMGRCAGVMHLFVVRAGVSRAMLLRLFLAM 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 307 EKGRIMEYECFYLYVYVYVAFRLPEPDGKGVEAFVADGELMVEAVOGVHBNYFMWYSGCV 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 OKGKHMELDCPYLVHVPVYVAFRLPEPDGKGVEAFVADGELMVEAVOGVHBNYFMWYSGCV 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 367 EPPPSWKQOKMPPEEP 383
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 DAPSGRDSRGRGPPPEEP 381
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: May 9, 2003, 16:54:55
 Job time : 49 secs

RESULT 14
AAV56053
ID AAV56053 standard; protein; 381 AA.
AC AAV56053;
XX 28-MAR-2000 (first entry)
DE Mouse sphingosine kinase 1a protein.
XX
XX Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
KM antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
KM cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
XX
OS Mus musculus.
XX
XX WO9961581-A2.
XX
XX 02-DEC-1999.
XX
XX 25-MAY-1999; 99WO-US11521.
XX
XX 26-MAY-1998; 98US-0086657.
XX 11-AUG-1998; 98US-0096049.
XX
XX (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.
XX
XX Spliegel S;
XX
XX WPI; 2000-072612/06.
XX N-PSDB; AAZ47167.
DR
XX
XX New sphingosine kinase, used to treat diseases involving abnormal cell
PT proliferation, e.g. cancer -
XX
XX Example 1; Fig 1; 116pp; English.
XX
XX This sequence represents the mouse sphingosine kinase 1a. The coding
CC sequence was isolated by searching a dbEST (expressed sequence tag
CC database) using amino acid sequence results from sequenced tryptic
CC digests of the rat sphingosine kinase. Expression of sphingosine kinase
CC in cells results in formation of sphingosine-1-phosphate, a known second
CC messenger, and confers serum-independent growth; increases proliferation,
CC and suppresses serum-deprivation or ceramide-induced apoptosis. The
CC sphingosine kinase nucleic acid is used: (a) to increase sphingosine
CC kinase content of cells, specifically for reducing cell death and/or
CC increasing cell proliferation; and (b) to produce transfected cells that
CC are used to screen for agents that inhibit or promote sphingosine kinase
CC activity. Agents that reduce sphingosine kinase activity or expression
CC are used: (i) to reduce cell proliferation, specifically for treating
CC cancer, and (ii) to treat diseases associated with abnormal cell
CC migration or motility, particularly cancer, restenosis or diabetic
CC neuropathy (but also atherosclerosis, stroke and Alzheimer's disease),
CC whereas agents that stimulate sphingosine kinase can be used to treat
CC conditions associated with reduced cell proliferation, e.g. developmental
CC retardation.
XX
XX Sequence 381 AA;
SQ
Query Match 80.7%; Score 1627.5; DB 21; Length 381;
Best Local Similarity 81.4%; Pred. No. 1.9e-167;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;
OY 7 PRCVLPKRCVLLVLPNGGKALQFRSHVQPLLAELISFTLMLTERNRHARELYRS 66
DB 6 PRCVLPKRCVLLVLPNGGKALQFRSHVQPLLAELISFTLMLTERNRHARELYRS 65
OY 67 FELGMDALVWMSGDGLMHEVYVNGIMBRPDWETATOKPLCSLPASSGNALASVNHVGY 126
DB 66 FELGMDALVWMSGDGLMHEVYVNGIMBRPDWETATOKPLCSLPASSGNALASVNHVGY 125

OY 127 EQVYNEDLLTNCCTLLCRLLSPNNLSLHTASGLRFLFSVLSLWMP1ADVLESEKYYR 186
DB 126 EQVYNEDLLTNCCTLLCRRLSPNNLSLHTASGLRFLFSVLSLWMP1ADVLESEKYYR 185
OY 187 LGBARFTLGTFLRLAALTFTYGRGLATLPVGRVGSCTPASPVPYVQGGPDALVPLEEVP 246
DB 186 LGEIRFTVGTFFRLASLRIRYOGQLATLPVGTVASRPAAS-TLVOKGPDTHLPLEEVP 244
OY 247 SHMTVPDEDFVLVALLSHLSGSEMFAPMGRCAQVMHLEFYRAGVSRAMLRLFLAM 306
DB 245 SHMTVPDEDFVLVALLSHLSGSEMFAPMGRCAQVMHLEFYRAGVSRAMLRLFLAM 304
OY 307 EGRHMEYECPLYVYPVAFRLPEKDGKGFVAVDGLMVEAVQGVHPYFMVWSCGV 366
DB 305 QKCKMELDCPLYVHVPVAFRLPEKDGKGFVAVDGLMVEAVQGVHPYFMVWSCGV 364
OY 367 EPPSWKPOQMPPEEP 383
DB 365 DAPSGRDSRGRPPPEEP 381
RESULT 15
ABB08087
ID ABB08087 standard; protein; 381 AA.
XX
XX ABB08087;
XX
XX 10-SEP-2002 (first entry)
XX
XX Murine sphingosine kinase 1a (mSPHK1a) protein sequence.
XX
XX Sphingosine kinase; SPHK; SPHK1a; cytosolic; vasotropic; antidiabetic;
XX neuroprotective; mouse; enzyme.
XX
XX Mus musculus.
XX
XX US2002042358-A1.
XX
XX 11-APR-2002.
XX
XX 02-MAR-2001; 2001US-0796487.
XX
XX 02-MAR-2000; 2000US-186352P.
XX
XX (SPIE/) SPIEGEL S.
XX
XX Spliegel S;
XX
XX WPI; 2002-478846/51.
XX
XX New isolated sphingosine kinase, useful in identifying modulators for
XX treating e.g. cancer, also related nucleic acid, vectors and
XX transformed cells -
XX
XX Disclosure: Fig 1; 24pp; English.
XX
XX The invention relates to an isolated sphingosine kinase (SPHK) DNA, cells
XX transformed with SPHK DNA are used to screen for agents that reduce,
XX eliminate or promote SPHK activity. Agents that inhibit activity are
XX useful for decreasing cell proliferation, e.g. for treating cancer, and
XX for treating diseases associated with abnormal migration and motility of
XX cells, e.g. restenosis or diabetic neuropathy. Agents that increase
XX activity are used to reduce cell death. Antibodies raised against SPHK,
XX and primers or oligonucleotides derived from the DNA are useful for
XX diagnosis. The antibodies are also useful as therapeutic inhibitors. The
XX present sequence represents a murine sphingosine kinase 1a (mSPHK1a)
XX corresponding to residues 124 to 504 of SPHK1a in GenBank Accession
XX no. AAC61597.
XX
XX Sequence 381 AA;
SQ
Query Match 80.7%; Score 1627.5; DB 23; Length 381;
Best Local Similarity 81.4%; Pred. No. 1.9e-167;

PI Spiegel S;
 XX
 DR WPI: 2000-072612/06.
 DR N-PSDB: AA247168.
 XX
 PT New sphingosine kinase, used to treat diseases involving abnormal cell
 proliferation, e.g. cancer -
 XX
 PS Example 1, Fig 1: 116pp: English.

CC This sequence represents the mouse sphingosine kinase 1a. The coding
 CC sequence was isolated by searching a dbEST (expressed sequence tag
 CC database) using amino acid sequence results from sequenced tryptic
 CC digests of the rat sphingosine kinase. Expression of sphingosine kinase
 CC in cells results in formation of sphingosine-1-phosphate, a known second
 CC messenger, and confers serum-independent growth; increases proliferation,
 CC and suppresses serum-deprivation or ceramide-induced apoptosis. The
 CC sphingosine kinase nucleic acid is used: (a) to increase sphingosine
 CC kinase content of cells, specifically for reducing cell death and/or
 CC increasing cell proliferation; and (b) to produce transfected cells that
 CC are used to screen for agents that inhibit or promote sphingosine kinase
 CC activity. Agents that reduce sphingosine kinase activity or expression
 CC are used: (i) to reduce cell proliferation, specifically for treating
 CC cancer, and (ii) to treat diseases associated with abnormal cell
 CC migration or motility, particularly cancer, restenosis or diabetic
 CC neuropathy (but also atherosclerosis, stroke and Alzheimer's disease),
 CC whereas agents that stimulate sphingosine kinase can be used to treat
 CC conditions associated with reduced cell proliferation, e.g. developmental
 CC retardation.

XX Sequence 388 AA:

Query Match 80.9%; Score 1630.5; DB 21; Length 388;
 Best Local Similarity 81.7%; Pred. No. 9,4e-168;
 Matches 308; Conservative 32; Mismatches 36; Indels 1; Gaps 1;

OY 7 PGCVLPKPCRVLYLVLPNPGCKGKALOLFESHVOPPLAEETSTMLTERNNHARELYRS 66
 DB 13 PGGLPFRPCRVLYLVLPNPGCKGKALOLFOSRVOPFLBEAETTKLITERNNHARELYCA 72
 OY 67 EELGMDALVYMSGDLMEHVNGLMERPDMTAIOKPLCSLPAGSGNALAASLNHYAGY 126
 DB 73 EELGMDALVYMSGDLMEHVNGLMERPDMTAIOKPLCSLPAGSGNALAASLNHYAGY 132
 OY 127 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESKYYR 186
 DB 133 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESKYYR 192
 OY 187 LGEHMFRTGLTLRLAALRTYRGLAYLPVGRVSKTPASPVYVOOGFVDAHLVPLEEPP 246
 DB 193 LGEHMFRTGLTLRLAALRTYRGLAYLPVGRVSKTPASPVYVOOGFVDAHLVPLEEPP 251
 OY 247 SHMTVYVPEODFVLYLVLTLSHSGEMFAAPMGCAAGVMHLFYVRAGYSRAALLRLFLAM 306
 DB 253 SHMTVYVPEODFVLYLVLTLSHSGEMFAAPMGCAAGVMHLFYVRAGYSRAALLRLFLAM 311
 OY 307 EKGHMEYECPTLYLVYVVAFAFLRPDKGVFAVDGELMSEAVOGGVHPNFMWVSCV 366
 DB 312 OKGHMELDCPTLYLVYVVAFAFLRPDKGVFAVDGELMSEAVOGGVHPNFMWVSCV 371
 OY 367 EPPSMKPOQMPPEEP 383
 DB 372 DAFSGRDSRGRPPPEEP 388

RESULT 13

AB08088 ID AB08088 standard; protein: 388 AA.

AC AB08088:

XX 10-SEP-2002 (first entry)

DE Murine sphingosine kinase 1b (msPHK1b) protein sequence.
 XX
 KM Sphingosine kinase; SPHK; SPHK1b; cytostatic; vasotropic; antidiabetic;
 KM neuroprotective; mouse; enzyme.
 XX
 OS Mus musculus.
 XX
 PN US2002042358-A1.
 XX
 PD 11-APR-2002.

PF 02-MAR-2001: 2001US-0796487.
 XX
 PR 02-MAR-2000: 2000US-186352P.

XX (SPIE/) SPIEGEL S.

PI Spiegel S;
 XX
 DR WPI: 2002-478846/51.

PT New isolated sphingosine kinase, useful in identifying modulators for
 PT treating e.g. cancer, also related nucleic acid, vectors and
 PT transformed cells -
 XX
 PS Disclosure; Fig 1: 24pp: English.

CC The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
 CC transformed with SPHK DNA are used to screen for agents that reduce,
 CC eliminate or promote SPHK activity. Agents that inhibit activity are
 CC useful for decreasing cell proliferation, e.g. for treating cancer, and
 CC for treating diseases associated with abnormal migration and motility of
 CC cells, e.g. restenosis or diabetic neuropathy. Agents that increase
 CC activity are used to reduce cell death. Antibodies raised against SPHK,
 CC and primers or oligonucleotides derived from the DNA are useful for
 CC diagnosis. The antibodies are also useful as therapeutic inhibitors. The
 CC present sequence represents a murine sphingosine kinase 1b (msPHK1b)
 CC corresponding to residues 1 to 388 of SPHK1b in Genbank Accession
 CC no. AAC61698.

XX Sequence 388 AA:

Query Match 80.9%; Score 1630.5; DB 23; Length 388;
 Best Local Similarity 81.7%; Pred. No. 9,4e-168;
 Matches 308; Conservative 32; Mismatches 36; Indels 1; Gaps 1;

OY 7 PGCVLPKPCRVLYLVLPNPGCKGKALOLFESHVOPPLAEETSTMLTERNNHARELYRS 66
 DB 13 PGGLPFRPCRVLYLVLPNPGCKGKALOLFOSRVOPFLBEAETTKLITERNNHARELYCA 72
 OY 67 EELGMDALVYMSGDLMEHVNGLMERPDMTAIOKPLCSLPAGSGNALAASLNHYAGY 126
 DB 73 EELGMDALVYMSGDLMEHVNGLMERPDMTAIOKPLCSLPAGSGNALAASLNHYAGY 132
 OY 127 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESKYYR 186
 DB 133 EOVNEDLLNCTLLCRRLSPMNLSTHTASGLRFSVLSLWGFIAVDLESKYYR 192
 OY 187 LGEHMFRTGLTLRLAALRTYRGLAYLPVGRVSKTPASPVYVOOGFVDAHLVPLEEPP 246
 DB 193 LGEHMFRTGLTLRLAALRTYRGLAYLPVGRVSKTPASPVYVOOGFVDAHLVPLEEPP 251
 OY 247 SHMTVYVPEODFVLYLVLTLSHSGEMFAAPMGCAAGVMHLFYVRAGYSRAALLRLFLAM 306
 DB 253 SHMTVYVPEODFVLYLVLTLSHSGEMFAAPMGCAAGVMHLFYVRAGYSRAALLRLFLAM 311
 OY 307 EKGHMEYECPTLYLVYVVAFAFLRPDKGVFAVDGELMSEAVOGGVHPNFMWVSCV 366
 DB 312 OKGHMELDCPTLYLVYVVAFAFLRPDKGVFAVDGELMSEAVOGGVHPNFMWVSCV 371
 OY 367 EPPSMKPOQMPPEEP 383
 DB 372 DAFSGRDSRGRPPPEEP 388

SQ Sequence 333 AA;
Query Match 86.9%; Score 1752; DB 23; Length 333;
Best Local Similarity 100.0%; Pred. No. 5e-181;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 MLTERRHNAELVSEELGKMDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAG 111
D 1 MLTERRHNAELVSEELGKMDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAG 60
QY 112 SGNAALASLNHYAGYEQVYTNEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVLSLAW 171
D 61 SGNAALASLNHYAGYEQVYTNEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVLSLAW 120
QY 172 GFIDVDVLESKRYRLGEMRFTGTFRLAALRTYRGRLAVLPGRGSKTPASPVVVOQ 231
D 121 GFIDVDVLESKRYRLGEMRFTGTFRLAALRTYRGRLAVLPGRGSKTPASPVVVOQ 180
QY 232 GPVDAHLVPLEEPYPSHMTVVPDEDFVLVLAHLHSHLGSEMFAPMGRCAAGVNHLEFVVR 291
D 181 GPVDAHLVPLEEPYPSHMTVVPDEDFVLVLAHLHSHLGSEMFAPMGRCAAGVNHLEFVVR 240
QY 292 AGVSRAMLRLFLAMEGRHMEYECPLYVYVYAFRLERKDGKGVRAVDGELMVSEAVQ 351
D 241 AGVSRAMLRLFLAMEGRHMEYECPLYVYVYAFRLERKDGKGVRAVDGELMVSEAVQ 300
QY 352 GOVHPNFMVYSGVEPPPSKPOOMPPPEPL 384
D 301 GOVHPNFMVYSGVEPPPSKPOOMPPPEPL 333

RESULT 11
ID AAE07883 standard; Protein: 382 AA.
AC AAE07883;
DT 01-NOV-2001 (first entry)
XX Mouse consensus sphingosine kinase (Sphk) protein.
DE Mouse; sphingosine kinase; Sphk; restenosis; ischaemia; gene therapy;
KW antisenesc therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW cytosolic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease.
OS Mus musculus.
PN MO200160990-A2.
XX 23-AUG-2001.
PD 14-FEB-2001; 2001WO-0504789.
PF 14-FEB-2000; 2000US-0182360.
PR 22-MAR-2000; 2000US-0191261.
XX (CUBA-) CURAGEN CORP.
PA (GETH) GENENTECH INC.
XX Rastelli L;
PI WPI: 2001-514770/56.
DR N-PSDB: AAD14425.
XX An isolated sphingosine kinase polypeptide useful for treating a
PT Sphk-associated disorder especially cancer, restenosis or ischemia in a
PT human.
PS Example 1: Page 92-94; 107pp; English.
XX The present invention relates to sphingosine kinase (sphk) polypeptides
CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-
CC

CC associated disorder especially cancers such as leukaemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin,
CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
CC or disorder associated with vascular diseases. Sphk gene is used in gene
CC therapy and antisenesc-therapy. Sphingolipids serving as signalling
CC molecules, have recently emerged as regulators of cell growth.
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC Sphk by tumor necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is consensus mouse sphingosine
CC kinase (Sphk) protein.
XX
SQ Sequence 382 AA;
Query Match 81.3%; Score 1639.5; DB 22; Length 382;
Best Local Similarity 80.9%; Pred. No. 9.7e-169;
Matches 310; Conservative 33; Mismatches 39; Indels 1; Gaps 1;

QY 1 MDPAGRGCVLPRECRVLYLNPFGGKKAQLFRSHVQPLAELISFTLMLTERRNA 60
D 1 MEPEECPRGLPRPCRVLYLNPFGGKKAQLFRSHVQPLAELISFTLMLTERRNA 60
QY 61 RELVSEELGKMDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSNAALASL 120
D 61 RELVCAEELGKMDALVYVSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSNAALASV 120
QY 121 NHVAGYEQVYTNEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVLSLAWGFIADVLE 180
D 121 NHVAGYEQVYTNEDLLTNCITLLCRRLSPMNLISLHTASGLRFSVLSLAWGFIADVLE 180
QY 181 SEKYRRLGEMRFTGTFRLAALRTYRGRLAVLPGRGSKTPASPVVVOGPDVAHLVP 240
D 181 SEKYRRLGEMRFTGTFRLAALRTYRGRLAVLPGRGSKTPASPVVVOGPDVAHLVP 239
QY 241 LEEPYPHMTVVPDEDFVLVLAHLHSHLGSEMFAPMGRCAAGVNHLEFVYRAGVSRAML 300
D 240 LEEPYPHMTVVPDEDFVLVLAHLHSHLGSEMFAPMGRCAAGVNHLEFVYRAGVSRALL 299
QY 301 RLFLAMEGRHMEYECPLYVYVYAFRLERKDGKGVFAVDGELMVSEAVQGOVHPNFM 360
D 300 RLFLAMQKGMHLEDCPLYVYVYAFRLERKDGKGVFAVDGELMVSEAVQGOVHPNFM 359
QY 361 MVSGCVPPEPPPSKPOOMPPPEPL 383
D 360 MVSGSRDAPSGDRSRGPPPEPL 382

RESULT 12
ID AAY56054 standard; Protein: 388 AA.
AC AAY56054;
DT 28-MAR-2000 (first entry)
XX Mouse sphingosine kinase 1b protein.
DE Mouse; sphingosine kinase 1b protein.
KW Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
KW antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
KW cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
OS Mus musculus.
PN WO9961581-A2.
XX 02-DEC-1999.
PD 25-MAY-1999; 99WO-0511521.
PF 26-MAY-1998; 98US-0086657.
PR 11-AUG-1998; 98US-0096049.
XX (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.
XX

A	E07882		ID AAE07882 standard; Protein; 384 AA.
X		AC	
X	AAE07882;	XX	
X		DT	01-NOV-2001 (first entry)
D		DE	Human sphingosine kinase (Sphk) protein #1.
O		KW	Human: sphingosine kinase: Sphk; restenosis; ischemia; gene therapy; cytotoxic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma; leukemia; vasculitic; cell proliferative disorder; vascular disease. Homo sapiens. MO200160990-AZ.
F		PX	23-AUG-2001.
X		PR	14-FEB-2001; 2001WO-US04789.
X		PP	14-FEB-2000; 2000US-0182360.
X		PT	22-MAR-2000; 2000US-0191261.
X	(CURA -) CURAGEN CORP.	PA	
X	(GETH) GENENTECH INC.	PI	Rastelli L;
X		DR	WP1: 2001-514770/56.
N	N-PDB: AMD14424.	PS	
A	An isolated sphingosine kinase polypeptide useful for treating a sphk-associated disorder especially cancer, restenosis or ischemia in a human	PT	
C	Claim 1: Page 90-91: 107pp: English.	XX	
C	The present invention relates to sphingosine kinase (SPHK) polypeptides and nucleic acids encoding them. SPHK is useful for treating a SPHK- associated disorder especially cancers such as leukemia, lymphoma, ovarian, breast, lung, colon, testicular, stomach and skin, atherosclerosis, restenosis or ischemia and cell proliferative disease or disorder associated with vascular diseases. SPHK gene is used in gene therapy and antisense-therapy. Sphingolipids serving as signalling molecules, have recently emerged as regulators of cell growth, differentiation, diverse cell phenotypes and cell death. Activation of SPHK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human endothelial cells. The present sequence is human sphingosine kinase (SPHK) protein.	CC	
SQ	Sequence 384 AA;	XX	
Query Match	97.3%; Score 1962; DB 22; Length 384;		
Best Local Similarity	97.7%; Pred: No. 1.le-203;		
Matches 375; Conservative	2; Mismatches 7; Indels		0;
OY	1 MPDAGCGPGCVLPKPCRCVLVLNPDGCKGAQLQFRSHVOPFLAEAFISTTLTERRNA	60	
Db	1 MPDAGCPGVLPKPCRCVLVLNPDGCKGAQLQFRSHVOPFLAEAFISTTLTERRNA	60	
OY	61 RELIVSEELGGMDLVVMGSDGIMHPYVNGLMERPDKETAIOKPLCSIPAGSGNALASL	120	
Db	61 RELIVNSEELGRMDLVVMGSDGIMHEPVYVNGLMERPMDETIOKPLCISIPAGSGNALASL	120	
OY	121 NHAYCEGYTMMEDLLTCNTLLCRRLISPMLSLAHASGLRIFSVLSIANGFIADVLE	180	
Db	121 NHAATGEYTMMEDLLTCNTLLCRRLPYSLPNULSLHTASSLSFSVLSIANGFIADVLE	180	
OY	181 SEKYRRLLCEAFRTLCFTLRALARTTYRGRLAYLPVGAVSGKTPPASPVVVVOGPVDAAHLVP	240	
Db	181 SDKYRRLLCEAFMTLTGTFLRLAIRTYRGRLATLPVGVGVRKTTPASPAPVVVOOCPVDAAHLVP	240	

QY	241	LEEP	PSHNTWVPPEDDEIVYIALLLSHSTSEMFAPMCKRCAGVMHLEFYRACVSRAML	300
DB	241	LEQVPSHNTWVPPEDDEIVYIALLLSHSTSEMFAPMCKRCAGVMHLEFYRACVSRAML	300	
QY	301	RLEFLAMEGRIMEFECYLYVYVVAFLRLEPPDKGCVAVANGELMSVANOGVHNRYM	360	
DB	301	RLEFLAMEGRIMEFECYLYVYVVAFLRLEPPDKGCVAVANGELMSVANOGVHNRYM	360	
QY	361	MVSGCEPPPSKRPQOMPPEEPL	384	
DB	361	MVSGCEPPPSKRPQOMPPEEPL	384	
RESULT 10				
ID	ABB90209	standard; Protein; 333 AA.		
AC	ABB90209;			
DT	24-MAY-2002	(first entry)		
DE	Human polypeptide SEQ ID NO 2585.			
KM	Cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;			
KM	antiallergic; hepatotropic; antidiabetic; antinflammatory; anticancer;			
KM	vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;			
KM	cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;			
KM	neurological disease; infection; human; secreted protein.			
OS	Homo sapiens.			
PN	WO200190304-A2.			
PD	29-NOV-2001.			
PF	18-MAY-2001; 2001MO-US16450.			
PR	19-MAY-2000; 2000US-205515P.			
PA	(HUMA-) HUMAN GENOME SCI INC.			
PI	Blise CE, Rosen CA;			
DR	WPI: 2002-122018/16.			
DR	N-PSDB: ABL90618.			
PT	Novel 1405 isolated polypeptides, useful for diagnosis, treatment and			
PT	prevention of neural, immune system, muscular, reproductive,			
PT	gastrointestinal, pulmonary, cardiovascular, renal and proliferative			
PT	disorders.			
XX	Claim 11: SEQ ID NO 2585; 2081bp + Sequence Listing; English.			
XX	The invention relates to novel genes (ABL89449-ABL90853) and proteins			
CC	(ABB89940-ABB90444) useful for preventing, treating or ameliorating			
CC	medical conditions e.g. by protein or gene therapy. The genes are			
CC	isolated from a range of human tissues disclosed in the specification.			
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful			
CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast			
CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone			
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;			
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune			
CC	hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's			
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative			
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias;			
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and			
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal			
CC	and parasitic infections.			
CC	Note: The sequence data for this patent did not form part of the			
CC	printed specification, but was obtained in electronic format directly			
CC	from WIPO at http://wipo.int/pub/published_pcl_sequences .			
XX				

44

DR WPI: 2002-47846/51.
XX
XX New isolated sphingosine kinase. useful in identifying modulators for
PT treating e.g. cancer, also related nucleic acid, vectors and
PT transformed cells -
XX
XX
PS Disclosure: Fig 3; 24pp; English.
XX
XX The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
CC transformed with SPHK DNA are used to screen for agents that reduce,
CC eliminate or promote SPHK activity. Agents that inhibit activity are
CC useful for decreasing cell proliferation, e.g. for treating cancer, and
CC for treating diseases associated with abnormal migration and motility of
CC cells, e.g. metastasis or diabetic neuropathy. Agents that increase
CC activity are used to reduce cell death. Antibodies raised against SPHK,
CC and primers or oligonucleotides derived from the DNA are useful for
CC diagnosis. The antibodies are also useful as therapeutic inhibitors. The
CC present sequence represents a human sphingosine kinase 1 (hSPHK1).
XX
XX Sequence 384 AA:
SQ
Query Match 99.6%; Score 2008; DB 23; Length 384;
Best Local Similarity 99.2%; Pred. No. 1.1e-208;
Matches 381; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 MDPAGGPRGVLPKRCRVLYLNPBGKGAQLFRSHVQPLLAEEISFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRVLYLNPBGKGAQLFRSHVQPLLAEEISFTLMTERRNHA 60
OY 61 RELVSEELGRMDALVYNSGDLMEHVNGLMERPDWETAIOKPLCSLPASSGNLAASL 120
DB 61 RELVSEELGRMDALVYNSGDLMEHVNGLMERPDWETAIOKPLCSLPASSGNLAASL 120
OY 121 NHVAGEVQVTEDELTNCTLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVDLE 180
DB 121 NHVAGEVQVTEDELTNCTLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVDLE 180
OY 181 SEKVRRLGEMRFTLTGFLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKVRRLGEMRFTLTGFLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPVPSTHTVPPDEDFVLYVALLSHLSGSEMFAPMGRCAGVNHLEFVVRAGVSRAAML 300
DB 241 LEEPVPSTHTVPPDEDFVLYVALLSHLSGSEMFAPMGRCAGVNHLEFVVRAGVSRAAML 300
OY 301 RFLAMEKGRHMEYECPTLVYVVAFAFLPEKDGKGFVAVDELAVSEAVOQVHPNFW 360
DB 301 RFLAMEKGRHMEYECPTLVYVVAFAFLPEKDGKGFVAVDELAVSEAVOQVHPNFW 360
OY 361 MVSGVEPPPSMKPQOMPPEEPL 384
DB 361 MVSGVEPPPSMKPQOMPPEEPL 384
RESULT 8
AA96057
ID AA96057 standard; Protein; 384 AA.
XX
XX AA96057;
AC
XX 05-DEC-2000 (first entry)
XX
XX Human sphingosine kinase A.
DE
XX Sphingosine kinase A; SKA; human; drug screening; infection;
XX antiinflammatory; antiallergic; anticancer; inflammation; allergy;
XX cancer; therapy; diagnosis.
OS Homo sapiens.
XX
XX
XX MO200052173-A2.
XX
XX 08-SEP-2000.
PD

XX
XX 02-MAR-2000; 2000W0-CA00223.
XX
XX
XX 02-MAR-1999; 99US-0122516.
XX
XX
XX (ALIX) NPS ALLELIX CORP.
XX
XX Munroe D, Gupta A, Falzone GR;
PI
XX
XX WPI: 2000-572185/53.
DR N-PSDB; AA50508.
XX
XX
XX The present sequence is that of human sphingosine kinase A (SKA),
CC an enzyme that phosphorylates sphingosine to form sphingosine
CC 1-phosphate. The sequence was deduced from that of a polynucleotide
CC (see AA50508) isolated from an HeLa cDNA library. The invention
CC provides polynucleotides (see AA50508-10) and polypeptides (see
CC AA96057-59) for the human sphingosine kinase (SK) homologues SKA,
CC SKB and SKC. The polypeptides can be obtained using recombinant
CC DNA methods, and host cells containing expression vectors including
CC SK polynucleotides are used in a claimed method of screening for
CC compounds that inhibit or activate human SK activity. Human SK
CC specific antibodies, inhibitors, ligands or their analogues can be
CC used as bioactive agents to treat inflammation or disease including
CC viral, bacterial or fungal infections, allergic responses,
CC mechanical injury associated with trauma, hereditary diseases,
CC lymphoma or carcinoma, and other conditions with activate the
CC genes of kidney, lung, heart, lymphoid or tissues of the nervous
CC system.
XX
XX Sequence 384 AA:
SQ
Query Match 99.2%; Score 2000; DB 21; Length 384;
Best Local Similarity 99.2%; Pred. No. 8e-208;
Matches 381; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 MDPAGGPRGVLPKRCRVLYLNPBGKGAQLFRSHVQPLLAEEISFTLMTERRNHA 60
DB 1 MDPAGGPRGVLPKRCRVLYLNPBGKGAQLFRSHVQPLLAEEISFTLMTERRNHA 60
OY 61 RELVSEELGRMDALVYNSGDLMEHVNGLMERPDWETAIOKPLCSLPASSGNLAASL 120
DB 61 RELVSEELGRMDALVYNSGDLMEHVNGLMERPDWETAIOKPLCSLPASSGNLAASL 120
OY 121 NHVAGEVQVTEDELTNCTLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVDLE 180
DB 121 NHVAGEVQVTEDELTNCTLLCRRLSPMNLISHTASGLRFSVLSLAWGFIADVDLE 180
OY 181 SEKVRRLGEMRFTLTGFLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
DB 181 SEKVRRLGEMRFTLTGFLRLAALRTYRGLAYLPVGRVSGKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPVPSTHTVPPDEDFVLYVALLSHLSGSEMFAPMGRCAGVNHLEFVVRAGVSRAAML 300
DB 241 LEEPVPSTHTVPPDEDFVLYVALLSHLSGSEMFAPMGRCAGVNHLEFVVRAGVSRAAML 300
OY 301 RFLAMEKGRHMEYECPTLVYVVAFAFLPEKDGKGFVAVDELAVSEAVOQVHPNFW 360
DB 301 RFLAMEKGRHMEYECPTLVYVVAFAFLPEKDGKGFVAVDELAVSEAVOQVHPNFW 360
OY 361 MVSGVEPPPSMKPQOMPPEEPL 384
DB 361 MVSGVEPPPSMKPQOMPPEEPL 384
RESULT 9

```

DB 121 NHVAGYEQVNTEDLLTNCITLLCRRLSPMNLSTHTASGLRFSVLSLWAGFIADVLE 180
OY 181 SEKYRRLCEMEFTLTGFTLRLLAALRTYRGRGLAYLPVGRVSKTPASPVVVOGQVDAHLVP 240
DB 181 SEKYRRLCEMEFTLTGFTLRLLAALRTYRGRGLAYLPVGRVSKTPASPVVVOGQVDAHLVP 240
OY 241 LEEPVPSHMTVVPDEDEVLVALLSHLCSSEMFAPMGRCAAGVNHFFYVAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDEVLVALLSHLCSSEMFAPMGRCAAGVNHFFYVAGVSRAML 300
OY 301 RFLAMEKGRHMEYECFYLIVVPAFRLEPKDGKGFVAVDGLMVSBAVGOVHPNYFW 360
DB 301 RFLAMEKGRHMEYECFYLIVVPAFRLEPKDGKGFVAVDGLMVSBAVGOVHPNYFW 360
OY 361 WVSGCVERPPPSMKPQOMPPEEPL 384
DB 361 WVSGCVERPPPSMKPQOMPPEEPL 384

RESULT 6
AAM40180
ID AAM40180 standard; protein; 384 AA.
XX
AC AAM40180;
XX
XX 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3325.
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QH, Zhou P, Goodrich R, Dirmacac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59336.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 5: SEQ ID NO 3325; 10078pp: English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AA42213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and

```

```

CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 384 AA:
XX
XX
Query Match 99.8%; Score 2011; DB 22; Length 384;
Best Local Similarity 99.7%; Pred. No. 5.1e-209;
Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MDPACGPRGVLPKCRVLYVLPNPGCGKALQIFRSHVQPLLAEEISFTLMTERRNHA 60
DB 1 MDPACGPRGVLPKCRVLYVLPNPGCGKALQIFRSHVQPLLAEEISFTLMTERRNHA 60
OY 61 RELVSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAPSL 120
DB 61 RELVSEELGRMDALVVMGSDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALAPSL 120
OY 121 NHVAGYEQVNTEDLLTNCITLLCRRLSPMNLSTHTASGLRFSVLSLWAGFIADVLE 180
DB 121 NHVAGYEQVNTEDLLTNCITLLCRRLSPMNLSTHTASGLRFSVLSLWAGFIADVLE 180
OY 181 SEKYRRLCEMEFTLTGFTLRLLAALRTYRGRGLAYLPVGRVSKTPASPVVVOGQVDAHLVP 240
DB 181 SEKYRRLCEMEFTLTGFTLRLLAALRTYRGRGLAYLPVGRVSKTPASPVVVOGQVDAHLVP 240
OY 241 LEEPVPSHMTVVPDEDEVLVALLSHLCSSEMFAPMGRCAAGVNHFFYVAGVSRAML 300
DB 241 LEEPVPSHMTVVPDEDEVLVALLSHLCSSEMFAPMGRCAAGVNHFFYVAGVSRAML 300
OY 301 RFLAMEKGRHMEYECFYLIVVPAFRLEPKDGKGFVAVDGLMVSBAVGOVHPNYFW 360
DB 301 RFLAMEKGRHMEYECFYLIVVPAFRLEPKDGKGFVAVDGLMVSBAVGOVHPNYFW 360
OY 361 WVSGCVERPPPSMKPQOMPPEEPL 384
DB 361 WVSGCVERPPPSMKPQOMPPEEPL 384

RESULT 7
AB808089
ID AB808089 standard; protein; 384 AA.
XX
AC AB808089;
XX
XX 10-SEP-2002 (first entry)
XX
DE Human sphingosine kinase 1 (hSPHK1) protein sequence.
XX
XX Sphingosine kinase; SPHK; SPHK1; cytosolic; vasotropic; antidiabetic;
XX neuroprotective; human; enzyme.
XX
OS Homo sapiens.
XX
XX
PN US2002042358-A1.
XX
PD 11-APR-2002.
XX
XX 02-MAR-2001; 2001US-0796487.
XX
XX 02-MAR-2000; 2000US-186352P.
XX
PA (SPIE/) SPIEGEL S.
XX
PI Spiegel S;
XX

```

KM T helper-1 related disease; chronic obstructive pulmonary disease;
 KM asthma; myocardial infarction; neurodegenerative disorder;
 KM wound healing; embryogenesis; anticoagulant; cerebroprotective;
 KM neuroprotective; antipruritic; antirheumatic; cytostatic; cardiant;
 KM vulnerable.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Region 16..95
 FT /label= Conserved_region
 PN WC200131029-A2.
 PD
 XX 03-MAY-2001.
 XX
 PF 27-OCT-2000; 2000MO-EP09498.
 XX
 XX 28-OCT-1999; 9905-0162307.
 PR 07-FEB-2000; 2000US-0180525.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PI Allen J, Gosink M, Melendez AJ, Takacs L;
 XX
 DR WPI: 2001-300510/31.
 DR N-PSDB; AAD04477.
 XX
 PT New human sphingosine kinase type I gene for screening drug candidates
 PT particularly inhibitors used for preventing or treating e.g.
 PT atherosclerosis, thrombosis, asthma and diabetes
 XX
 PS Claim 3; Fig 1; 91pp; English.
 XX
 CC The present sequence is human sphingosine kinase type I (hsk1).
 CC The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
 CC substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
 CC and encoded polypeptide are applicable in screening drug candidates
 CC particularly inhibitors for preventing or treating disorders such as
 CC haemostasis, thrombosis, allergic reactions, proliferative diseases
 CC including cancer, haematopoietic disorders such as leukemia,
 CC cardiovascular diseases such as stroke, atherosclerosis and coronary
 CC artery disease, dyslipidaemia, diabetes including type I and type II
 CC diabetes, autoimmune and inflammatory diseases such as multiple
 CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
 CC disease, asthma, myocardial infarction, neurodegenerative disorders,
 CC natural wound healing processes and embryogenesis.
 CC
 XX
 SQ Sequence 384 AA;
 Query Match 100.0%; Score 2016; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.5e-209;
 Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPAGGPRGVLPKRCRVYVLLNPRGKGKALQFRSHVOPPLAAEISFTLMTERRNHA 60
 DB 1 MDPAGGPRGVLPKRCRVYVLLNPRGKGKALQFRSHVOPPLAAEISFTLMTERRNHA 60
 QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIOKPLCSIPAGSGNALAASL 120
 DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIOKPLCSIPAGSGNALAASL 120
 QY 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISHTASGRLFSVSLAMGFIADVDLE 180
 DB 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISHTASGRLFSVSLAMGFIADVDLE 180
 QY 181 SEKYRRLGEMFTTIGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
 DB 181 SEKYRRLGEMFTTIGTFLRLAALRTYRGRLAYLPVGRVSKTPASPVVVOGQPVDAHLVP 240
 QY 241 LEEPPSHWTVVPPDEDFVLVALLSHLSGSEMFAPKGRCAAGVNHLLFYVAGVSRAML 300
 DB 241 LEEPPSHWTVVPPDEDFVLVALLSHLSGSEMFAPKGRCAAGVNHLLFYVAGVSRAML 300

QY 301 RLFLAMEKGRHMEYECYLVVPPVAFRLPEKDGKGYFAVDGELMSEAVOGVHPNFW 360
 DB 301 RLFLAMEKGRHMEYECYLVVPPVAFRLPEKDGKGYFAVDGELMSEAVOGVHPNFW 360
 QY 361 MVSGVPEPPSWKPKQMPPEEPL 384
 DB 361 MVSGVPEPPSWKPKQMPPEEPL 384
 RESULT 5
 AAB48007
 ID AAB48007 standard; Protein: 384 AA.
 AC AAB48007;
 DT 19-MAR-2001 (first entry)
 DE Human sphingosine kinase protein sequence.
 DE
 XX Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
 KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 KW human.
 XX Homo sapiens.
 OS
 XX
 PN WO200070028-A1.
 XX
 XX 23-NOV-2000.
 PD
 PF 12-MAY-2000; 2000MO-AV00457.
 XX
 PR 13-MAY-1999; 99AU-0000339.
 PR 08-JUL-1999; 99AU-0001504.
 XX
 PA (JOHU) JOHNSON & JOHNSON RES PTY LTD.
 PI Pltson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vades MA;
 DR N-PSDB; AAC84161.
 DR
 XX Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock
 XX
 XX
 PS Claim 9; Fig 7a; 100pp; English.
 XX
 CC This represents a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.
 CC
 XX
 SQ Sequence 384 AA;
 Query Match 99.9%; Score 2013; DB 22; Length 384;
 Best Local Similarity 99.7%; Pred. No. 3.1e-209;
 Matches 383; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPAGGPRGVLPKRCRVYVLLNPRGKGKALQFRSHVOPPLAAEISFTLMTERRNHA 60
 DB 1 MDPAGGPRGVLPKRCRVYVLLNPRGKGKALQFRSHVOPPLAAEISFTLMTERRNHA 60
 QY 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIOKPLCSIPAGSGNALAASL 120
 DB 61 RELVSEELGRMDALVWMSGDLMEHVNGLMERPDMETAIOKPLCSIPAGSGNALAASL 120
 QY 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISHTASGRLFSVSLAMGFIADVDLE 180
 DB 121 NHVAGYEDVNTEDLLTNCITLLCRRLSPMNLISHTASGRLFSVSLAMGFIADVDLE 180

Best Local Similarity 100.0%; Pred. No. 1.5e-209;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGRCVLPFRPCRYVLYLNPGRGKALQFRSHVOPLLAEISFTMLTERRNA 60
Db 1 MDPAGGRCVLPFRPCRYVLYLNPGRGKALQFRSHVOPLLAEISFTMLTERRNA 60
QY 61 RELVSEELGEMDLYVNSGDLMEHVNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
Db 61 RELVSEELGEMDLYVNSGDLMEHVNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
QY 121 NHYAGYEOVTNEDLLTNCITLLICRLLSPMNLISLHTASGLRLESVLSLAMEFIADVLE 180
Db 121 NHYAGYEOVTNEDLLTNCITLLICRLLSPMNLISLHTASGLRLESVLSLAMEFIADVLE 180
QY 181 SEKRYRLGEMFTLGTEFLRLAALRTYRGRLAYLPVGRVSKTPASPVVQOGFVDANLVP 240
Db 181 SEKRYRLGEMFTLGTEFLRLAALRTYRGRLAYLPVGRVSKTPASPVVQOGFVDANLVP 240
QY 241 LEEPVSHMTVVPDEDFVLYLALSHSLGSEMPAARPGCAAGVMHLFVYRAGVSRAML 300
Db 241 LEEPVSHMTVVPDEDFVLYLALSHSLGSEMPAARPGCAAGVMHLFVYRAGVSRAML 300
QY 301 RLFLAMEGRHMEYECPLYLVYVPAFLPEKDGKVPADVDELMSAVOQGVHNFYFW 360
Db 301 RLFLAMEGRHMEYECPLYLVYVPAFLPEKDGKVPADVDELMSAVOQGVHNFYFW 360
QY 361 MWSGCVPEPPPSMKPOOMPPEEPL 384
Db 361 MWSGCVPEPPPSMKPOOMPPEEPL 384

RESULT 3
AAB94589 standard; Protein: 384 AA.

QY 26-JUN-2001 (first entry)
Human protein sequence SEQ ID NO:15394.
Human: primer: detection; diagnosis; antisense therapy; gene therapy.
OS Homo sapiens.
PN EP1074617-A2.
PD 07-FEB-2001.
PE 28-JUL-2000; 2000EP-0116126.
PF 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 8: SEQ ID 15394; 2537pp + CD ROM, English.
XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to a 5'-end
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the combination of
CC oligonucleotide and sequence is selected from those defined in
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers are obtained by the method of AAH13632
CC and AAH13633 to AAH1842, represent human cDNA sequences: AAH92446 to
CC AAH93893, represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

Sequence 384 AA:
SQ

Query Match 100.0%; Score 2016; DB 22; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.5e-209;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAGGRCVLPFRPCRYVLYLNPGRGKALQFRSHVOPLLAEISFTMLTERRNA 60
Db 1 MDPAGGRCVLPFRPCRYVLYLNPGRGKALQFRSHVOPLLAEISFTMLTERRNA 60
QY 61 RELVSEELGEMDLYVNSGDLMEHVNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
Db 61 RELVSEELGEMDLYVNSGDLMEHVNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
QY 121 NHYAGYEOVTNEDLLTNCITLLICRLLSPMNLISLHTASGLRLESVLSLAMEFIADVLE 180
Db 121 NHYAGYEOVTNEDLLTNCITLLICRLLSPMNLISLHTASGLRLESVLSLAMEFIADVLE 180
QY 181 SEKRYRLGEMFTLGTEFLRLAALRTYRGRLAYLPVGRVSKTPASPVVQOGFVDANLVP 240
Db 181 SEKRYRLGEMFTLGTEFLRLAALRTYRGRLAYLPVGRVSKTPASPVVQOGFVDANLVP 240
QY 241 LEEPVSHMTVVPDEDFVLYLALSHSLGSEMPAARPGCAAGVMHLFVYRAGVSRAML 300
Db 241 LEEPVSHMTVVPDEDFVLYLALSHSLGSEMPAARPGCAAGVMHLFVYRAGVSRAML 300
QY 301 RLFLAMEGRHMEYECPLYLVYVPAFLPEKDGKVPADVDELMSAVOQGVHNFYFW 360
Db 301 RLFLAMEGRHMEYECPLYLVYVPAFLPEKDGKVPADVDELMSAVOQGVHNFYFW 360
QY 361 MWSGCVPEPPPSMKPOOMPPEEPL 384
Db 361 MWSGCVPEPPPSMKPOOMPPEEPL 384

RESULT 4
AAB00924 standard; Protein: 384 AA.

QY 04-JUL-2001 (first entry)
Human sphingosine kinase type 1 (hsk1).
DE Human: sphingosine kinase type 1; sk1; chromosome 17q25.2;
XX sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
XX thrombosis; allergic reaction; proliferative disease; cancer;
KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KW autoimmune disease; inflammatory disease; multiple sclerosis;

QY 345 MYSEAV 350
 Db 334 VLKRAV 339

RESULT 11
 SM4B_MOUSE
 ID SM4B_MOUSE STANDARD: PRT: 782 AA.
 AC 062179:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
 GN SEMA4B OR SEMAC OR SEMC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRI; TISSUE=Brain;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Puschel A.W., Adams R.H., Betz H.;
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-946(1995).
 RN [2]
 RP INTERACTION WITH GIPC.
 RX MEDLINE=99253973; PubMed=10318831;
 RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
 RT "A PDZ protein regulates the distribution of the transmembrane
 semaphorin, M-SemF.";
 RL J. Biol. Chem. 274:14137-14146(1999)
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
 CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
 LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
 MODERATE LEVELS FROM THEN UNTIL BIRTH.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X85992; CAA59984.1; .
 DR MGD: MGI:107559; Sema4b.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01403; Sema; 1.
 DR Pfam: PF01437; PSI; 1.
 DR SMART: SM00423; PSI; 1.
 KM Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
 KV Developmental protein; Glycoprotein.
 FT NON_TER 1
 FT DOMAIN 1
 FT TRASNEM 662
 FT DOMAIN 663
 FT DOMAIN 683
 FT DOMAIN 684
 FT DOMAIN 782
 FT DOMAIN 782
 FT DOMAIN 782
 FT DOMAIN 782
 FT DOMAIN 782
 FT DOMAIN 782
 FT DOMAIN 782
 FT CAROHD 12
 FT CAROHD 12
 FT CAROHD 15

FT CAROHD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 574 574 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 782 AA; 86823 MW; 627A81FC8F87AC8 CRC64;

Query Match 4.6%; Score 92; DB 1; Length 782;
 Best Local Similarity 19.1%; Pred. No. 41; Indels 138; Gaps 19;
 Matches 81; Conservative 49; Mismatches 156;

QY 29 KALQFRSHVQPLAEAEISFTLTERNRHARELYRSEELG---RMDALVYMSGDLMH 85
 Db 364 RVLNFLKDH---FLNDGVRSKLLQLPPARKQVAVHVRPLGHSYVDVLFCTGGRH 420
 QY 86 EVVNGIMERPDMETAIO-----KPLCSLPAGSGMALASINHYAGEQVYTNEDLTNCTL 140
 Db 421 KAVT-LSSRVHIIIELOIFPGQPVQNLDSHGILYASSH-SGVQVP---VANCSL 474
 QY 141 L-LCRRLSPMLLSHTSGLRFS-----VLSLWAGFIADVDSEKTRRIGEMRTL 194
 Db 475 YPCGDCLARDPYCAWTSACRLASLYOPDLASRPW--TDEGASVK----- 521
 QY 195 GTEFLRALRTYRGRLAYLPVGRVSKTPASPVYVOOGPVDAHLVPLEPVPSH-WT--- 250
 Db 522 ----ELCKNSSTKAR--FLVPGK-----FKQYQIQPNTVNTIACLLSLNLRVHNG 570
 QY 251 -----VPEDEDFVLVLLSHLSGSEMPAPMGCAQVMHLYVYRAGVSAMLL 300
 Db 571 APVNASAGRVLPDGLLV-----GSOQ-----GLGVFCWSTIEEGFOOLVAS 614
 QY 301 RLFLAMEK-----RHMVECPVLYVP 323
 Db 615 YCEVMEEGMDQKRDGTPIIINTSRVASAPAGRDSMGADKSYNNEELVMTLTFVPM 674
 QY 324 VVAFR---EPKKGCVFAVDCELMTSEAVOGVHNFMVSGVPEPPSMKPP 380
 Db 675 VLFLEFLYRHRGMLFLKQGE-----CASVHKTRPYLVLP 713

QY 381 EEPL 384
 Db 714 TRPL 717

RESULT 12
 NAL1_HUMAN
 ID NAL1_HUMAN STANDARD: PRT: 1473 AA.
 AC 09C000: Q9Y2E0; Q9HAV8; Q9BZ29; Q9BZ28; Q9UFT4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-
 forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., Distefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 proteins.";
 RL Cell Death Differ. 7:1273-1274(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martiny-Bar F., Hofmann K., Tschoopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family

RT Implicated in apoptosis and inflammation.";
 CC Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Erythrocytes.
 RX MEDLINE-21153743; PubMed-11076957;
 RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
 RA Shi M., Vincenz C., Ward P.A.;
 RT Molecular cloning and characterization of DEFCAP-L and -S, two
 RT isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 RT proteins.";
 RL J. Biol. Chem. 276:9230-9238(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.
 RC TISSUE-T-cell;
 RX MEDLINE-21153744; PubMed-1113115;
 RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewski S.,
 RA Godzik A., Reed J.C.;
 RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
 RT c-dependent caspase activation and apoptosis.";
 RL J. Biol. Chem. 276:9239-9245(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Brain;
 RX MEDLINE-99246063; PubMed-10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohta O.;
 RT Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE-Uterus;
 RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wleemann S.;
 RL Submitted (Sep-1999) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Able to form cytoplasmic structures termed death
 CC effector filaments. Enhances Apaf1 and cytochrome c-dependent
 CC activation of pro-caspase-9 and consecutive apoptosis.
 CC Seems to bind ATP.
 CC SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 CC and with Apaf1 in a cytochrome c-inducible way leading to the
 CC formation of an apoptosome. This interaction may be ATP-dependent.
 CC SUBCELLULAR LOCATION: Cytoplasmic.
 CC - ALTERNATIVE PRODUCTS: 4 isoforms: 1/NAC beta/DEFCAP-L (shown
 CC here), 2/NAC alpha/DEFCAP-S, 3/NAC gamma and 4/NAC delta; are
 CC produced by alternative splicing
 CC TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 CC expressed in peripheral blood leukocytes, chronic myelogenous
 CC leukemia cell line K-562, followed by thymus, spleen and heart.
 CC Also detected in lung, placenta, small intestine, colon, kidney,
 CC liver and muscle.
 CC SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 CC SIMILARITY: CONTAINS 1 NACHT DOMAIN.
 CC SIMILARITY: CONTAINS 1 CARD DOMAIN.
 CC - SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR HSP: P13489; 1AAY.
 DR MIM: 606636;
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003590; LRR_RNinh.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF00560; LRR_1.
 DR Pfam: PF02758; PAAD_DAPIN_1.
 DR PRINTS: PR00364; DISEASERISIT.
 DR SMART: SM00368; LRR_R1; 5.
 DR PROSITE: PS50209; CARD; 1.
 DR PROSITE: PS50824; DAPIN; 1.
 DR Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 92
 FT REPEAT 328 637 DAPIN.
 FT REPEAT 704 725 NACHT.
 FT REPEAT 807 830 LRR 1.
 FT REPEAT 864 887 LRR 2.
 FT REPEAT 921 944 LRR 3.
 FT REPEAT 950 973 LRR 4.
 FT REPEAT 1199 1215 LRR 5.
 FT REPEAT 1216 1236 LRR 6.
 FT DOMAIN 1374 1463 LRR 7.
 FT NF_BIND 354 341 CARD.
 FT VARSPLIC 958 987 MISSING (IN ISOFORM 3 AND ISOFORM 4).
 FT VARSPLIC 1262 1305 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT MUTAGEN 340 340 K->L: NO EFFECT.
 FT MUTAGEN 340 340 K->S: NO EFFECT.
 FT CONFLICT 155 155 L->H (IN REF. 1).
 FT CONFLICT 246 246 T->S (IN REF. 1).
 FT CONFLICT 782 782 T->S (IN REF. 1).
 FT CONFLICT 878 878 T->M (IN REF. 1).
 FT CONFLICT 995 995 T->I (IN REF. 1).
 FT CONFLICT 1119 1119 M->V (IN REF. 1).
 FT CONFLICT 1184 1184 M->V (IN REF. 1).
 FT CONFLICT 1241 1241 V->L (IN REF. 1 AND 6).
 FT CONFLICT 1366 1366 R->C (IN REF. 1).
 SO SEQUENCE 1473 AA; 165865 MW; 438F0DC45C2562D CRC64;
 Query Match 4.68; Score 92; DB 1; Length 1473;
 Best local Similarity 23.28; Pred No. 9.4; Mismatches 107; Indels 146; Gaps 21;
 Matches 89; Conservative 42;
 Db 78 MSCDGLMEHVVGMLERDMEFALOKFLCSLP-----AGSG-----NMLASLNHYAGYE 127
 Db 815 LSGNSLSHSAVAKSLK-----TLRRPCLLETLRAGGGLTAEDCKDLAFGLR--ANQ 865
 Db 128 QVTNEDLTLNC-----TLLICRRLSP-----MNLISLHTASGL--RLSVSLAMGFT 174
 Db 866 TLTELDLSFNVLTDGAKANHLQRLQPSCKLQRLQLVSCGLTSDCCDLASVLS-ASPSL 924
 Db 175 ADVNDESE-----KYRRLEGMRFTLGFLR--LAALRTYNGRL--- 210
 Db 925 KELDLONNLDGVGRLLCEGLRHPACKLIRGLIDQTLTSDMRQELRALEQEKPOLITF 984
 Db 211 -----AYLVGRVSGSKTPASPVVVOQGVPAHLVPL----- 241
 Db 985 SRKPSVMPPTGGLDTGEMSNSTSLKRLQRLGSEKRAVSH-----AOANIKLIDVSKIP 1038
 Db 242 -----EPPVPHMTVVVDEDFVYLAL-----LISH-LGS-EWFAAPMGRCAGVW 285
 Db 1039 PLAEIAEESPE--VVVE-----LDCVSPASQGLIRKPLGTDDDFWGPVAVEV- 1090
 Db 286 HLFYVAGVSRAMLLRLFLAMEGRHNEYCPYLVVPAVFRLEPRDG-----KGFVAV 340
 Db 1091 -----VDKKNL-----IRVHPVAGSTRMPNTGLGCFVWREAVTV 1125
 Db 341 DGLVASEAVQGVHPNFWVSG 364
 Db 1126 EIEFCVMDQFLGEINPOHSMWVAG 1149

RESULT 13
PR66_CAEEL STANDARD: PRT: 414 AA.
AC P46502:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable 26S protease regulatory subunit 6B.
GN F23F12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du 2.1;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
CC DEGRADATION OF UBQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
CC COMPLEX CONTERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC 26S COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U12965; AAA20608.1; -
DR Wormpep; F23F12.6; CE01253.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_centre.
DR InterPro: IPR003960; AAA_sub.
DR Pfam: PF00004; AAA; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRFAMs: TIGR01242; 26Sp45; 1.
DR PROSITE: PS00674; AAA; 1.
KW Hypothetical protein; Proteasome; ATP-binding; Nuclear protein.
FT NP_BIND 202 209 ATP (POTENTIAL).
SQ SEQUENCE 414 AA; 46358 MM; 8630AFPBBAC2CF32 CRC64;
Query Match 4.48; Score 89; DB 1; Length 414;
Best Local Similarity 22.08; Pred. No. 3.2;
Matches 46; Conservative 27; Mismatches 78; Indels 58; Gaps 5;
OY 42 LAEAFSTLTALTEERRNARELVSEELGRMDALVVMGSDGLMEHYVNGLMERPMEFAI 101
DB 47 LAHLOVMEDYIKLETRNLEKELHAQE-----EV----- 75
OY 102 QKPLCSLPAGSGNALAASLNHYAGYQVNTEDLTNCTLLCRLLSPMNLSTHTASGL 161
DB 76 -KRIGSVPLVIGQLEAVDQNHAIYGTGSNYRVRLSILDRLLKRGCSVALHKYEN- 133
OY 162 RLFSVL-----SLAWGFADVDLESEKRYRLGEMRTTLGTFLELAALR 204
DB 134 ALVDVLPREADSIQMLRPDEKPDISTGYDGLDMOKOEVEAVELPLTHGELYOOIGID 193
OY 205 TYRGRIATLPVG-----RVGSKTPAS 225
DB 194 PPRGVLMYGPCCGKTMAKAVAAVTAAAS 222
RESULT 14
SERA_RAT STANDARD: PRT: 533 AA.
AC O08651:
DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.1.95) (3-PGDH).
GN PGDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250633; PubMed=9163325;
RA Achouri Y., Rider M.H., van Schaftingen E., Robbi M.;
RT "Cloning, sequencing and expression of rat liver 3-phosphoglycerate
RT dehydrogenase.";
RL Biochem J. 323:365-370(1997).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-
CC phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, BRAIN, TESTIS.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X97772; CAA66374.1; -
DR HSSP: P08328; 1PSD.
DR InterPro: IPR002162; D_2hyd-ac-dh.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH_C; 1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
KW Serine biosynthesis; Oxidoreductase; NAD.
FT ACT_SITE 236 236 SUBSTRATE-BINDING (BY SIMILARITY).
FT ACT_SITE 265 265 BY SIMILARITY.
FT ACT_SITE 283 283 BY SIMILARITY.
SQ SEQUENCE 533 AA; 56493 MM; 7273DAC3349E95EP CRC64;
Query Match 4.48; Score 89; DB 1; Length 533;
Best Local Similarity 22.78; Pred. No. 4.3;
Matches 68; Conservative 42; Mismatches 77; Indels 112; Gaps 17;
OY 51 LMTERRNINARELVSEELGRMDALVVMGSDGLMEHYVNGL-----MERPME 98
DB 28 LQVNRKQNLKRELLA-ELQDCEGLIVASATVYADVINAARLQVYGRAGCVNDVLE 86
OY 99 TAIQPLCSLPAGSGNALAASLNHYAGYQVNTEDLTNCTLLCRLLSPMNLSTHTA 158
DB 87 AATRGVILVMNPNPNCNSLSAA-----ELT-----CGMLMCIARQIDQ----- 123
OY 159 SGLRLFSVLSTAWGFIADVDLESEKRYRLGEMRTTLGTFLELAALRYRGR-LATLPVGR 217
DB 124 -----ATASMKDKWDR-----KKFMGTFL-----NGKTLGILGLGR 155
OY 218 VGSKTPA-----SPVV-----VQGPDAHLVPLEEPVSHMTVAPDED 256
DB 156 IGRVAAARMQAMGKMTVGVDPTIISPEVAASPVQO-----LPLFE-----IWLPLCD 201
OY 257 FVLV-LALHSHLG---SEMPAAPMGRCAAGYMLFYVYAG-VSRAMLRLFLAMEKGR 310
DB 202 FITVHTPLPSTTGILNDSTFA---QCKKGVYVNCARGIVDEGALLR---ALQSCG 253
RESULT 15
SERA_HUMAN STANDARD: PRT: 533 AA.
ID SERA_HUMAN

AC	043175; Q9B001; .
AD	15-JUL-1999 (Rel. 38, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
GN	PGDH OR PGDH3.
OS	Homo sapiens (human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxId=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Medline-20178699; PubMed-10713460.
RX	Cho H.M., Jun D.Y., Bae M.A., Ahn J.D., Kim Y.H.;
CC	"Nucleotide sequence and differential expression of the human
RT	3-phosphoglycerate dehydrogenase gene.";
RL	Gene 245:193-201(2000).
RN	[2]
RP	SEQUENCE FROM N.A. AND VARIANTS PGDH DEFICIENCY MET-425 AND MET-490.
RA	Medline-20530221; PubMed-11058955;
RX	Klomp L.W.J., de Koning T.J., Mallinger H.E.M., van Beurden E.A.C.M.,
CC	Bink M., Opm A.F.L., Duran M., Jaeken J., Plineda M.,
CC	van Maldergem I., Poll-The B.T., van den Berg I.E.T., Berger R.;
CC	"Molecular characterization of 3-phosphoglycerate dehydrogenase
CC	deficiency -- a neurometabolic disorder associated with reduced
CC	L-serine biosynthesis.";
CC	Am. J. Hum. Genet. 67:1389-1399(2000).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Tissue-brain, lung, and muscle:
CC	Strausberg R.;
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC	-1 CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC	phosphohydroxypropylate + NADH.
CC	-1 PATHWAY: Serine biosynthesis: first step.
CC	-1 SUBUNIT: Homotrimer (By similarity).
CC	-1 Disease: Defects in PGHD are the cause of a deficiency
CC	characterized by congenital microcephaly, psychomotor retardation,
CC	and seizures.
CC	-- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC	DEHYDROGENASES FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AF006043; AAB88664.1; .
DR	EMBL; AF171237; AAD51415.1; .
DR	EMBL; BC000303; AAH00303.1; .
DR	EMBL; BC001349; AAH01349.1; .
DR	EMBL; BC011262; AAH11262.1; .
DR	HSSP; P08328; 1PSD.
DR	GeneW; HGNC:8923; PGDH.
DR	MIM; 606879; .
DR	MIM; 601815; .
DR	InterPro; IPR002162; D_2Hydc_dh.
DR	Pfam; PF00389; 2-Hcid_Dh_C.1
DR	Pfam; PF02826; 2-Hcid_Dh_C.1
DR	PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR	PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR	PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
FM	Serine biosynthesis; Oxidoreductase; NMD; Disease mutation.
FT	ACT_SITE 236 .. 236
FT	ACT_SITE 265 .. 265
FT	ACT_SITE 283 .. 283
FT	VARIANT 425 .. 425
FT	BY SIMILARITY.
FT	V->M (IN PGDH DEFICIENCY).
FT	/FTID-VAR_013461.
FT	V->M (IN PGDH DEFICIENCY).
FT	/FTID-VAR_013462.
FT	VARIANT 490 .. 490
FT	V->M (IN PGDH DEFICIENCY).
FT	/FTID-VAR_013462.

FT	CONFLICT	25	D ->E. (1M REF. 1.)
SQ	SEQUENCE	533 AA: 56650 MW: C58EB72275C45B35 C8C64;	
	Query Match	4.3%: Score 87; DB 1; Length 533;	
	Best Local Similarity	23.1%: Pred. No 6.7;	
	Matches	69; Conservative .42; Mismatches 76; Indels 112; Gaps 18;	
QY	51	LMATERRRHARELYRSSEELGRMDALVYNSGGGLMHEVYNGL-----MERDME 98	
Db	28	LVKVERKQULSSEELIA-ELDDEGLIVSAKRYADVINAERKLOYVGRAGTGVNDVILE 86	
QY	99	TAIOKPLCCSPAGSGNALASLNHYAGYEQVNTEDLLTNCSTLCRLLSPNNLLSLHTA 158	
Db	87	AATFRKGILLVMTPTNGNSLSAA-----ELT-----CGMTCLABQIDPQ-----ATA 126	
QY	159	SGLRILFSVLSLAWGFIADVDLSEKYYRRIIGEMRPTLQTLPLLAALRYRGR-LAYLPLVR 217	
Db	127	S-----MKDGKMERKKF-----MGTEL-----NGRTGLILDLGR 155	
QY	218	VGSKTPA-----SPVY-----VOOGVGDHNLVPLDEEPPVSHMTIVVDEED 256	
Db	156	IGREAVATMDSFGMKITGYDPIILSPYASNSGYQO-----LPLEE-----INPLCD 201	
QY	257	FVLV-LALLSHSLG---SEMFAPMGRCACAGVMLLFYRAG-VSRAMLLRLFLAMEGR 310	
Db	202	FLTYVITPLPSTGGLINDNTFA---QCKKGVRYVNCARGIVDGBQALR---ALQSGQ 253	

Search completed: May 9, 2003, 16:55:19
Job time : 19 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:53:49 ; Search time 22 Seconds
(without alignments)
1677.983 Million cell updates/sec

Title: US-09-937-060a-5

Perfect score: 2016

Sequence: 1 MDPAGGPGVLPFRCHVLV.....CVPPEPSMKPQMPPEPL 384

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

PIR_73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451	22.4	1240	2 T05162	hypothetical prote
2	374.5	18.6	473	2 T19707	hypothetical prote
3	356.5	17.7	687	2 S51398	hypothetical prote
4	355	17.6	458	2 T38776	hypothetical prote
5	345	17.1	624	2 S67059	hypothetical prote
6	250.5	12.4	549	2 T33517	hypothetical prote
7	155.5	7.7	306	2 AH1769	conserved hypotet
8	140.5	7.0	306	2 AE1394	conserved hypotet
9	136.5	6.8	309	2 H69995	hypothetical prote
10	130	6.4	295	2 A83894	hypothetical prote
11	129.5	6.4	303	2 F69795	conserved hypotet
12	127	6.3	342	2 D86677	transcription regu
13	126	6.2	302	2 F86849	transcription regu
14	124.5	6.2	310	2 AG1665	hypothetical prote
15	122.5	6.1	311	2 F83871	multidrug resistan
16	122.5	6.1	364	2 F84898	hypothetical prote
17	119.5	5.9	310	2 A11293	hypothetical prote
18	118	5.9	309	2 AH1528	conserved hypotet
19	117	5.8	315	2 AB2166	hypothetical prote
20	113	5.6	309	2 AF1171	conserved hypotet
21	111	5.5	294	2 G95120	conserved hypotet
22	111	5.5	309	2 H70861	hypothetical prote
23	111	5.5	311	2 C97990	conserved hypotet
24	111	5.5	321	2 C70596	hypothetical prote
25	106	5.3	315	2 A89978	conserved hypotet
26	105.5	5.2	295	2 D83734	hypothetical prote
27	104	5.2	304	2 F72386	conserved hypotet
28	103	5.1	297	2 F69595	multidrug resistan
29	103	5.1	1273	2 E72611	probable ATP-depen

30	100.5	5.0	433	2 S75948	hypothetical prote
31	99	4.9	449	2 C83634	hypothetical prote
32	97.5	4.8	305	2 C89844	hypothetical prote
33	97	4.8	311	2 D75405	conserved hypotet
34	94	4.7	899	2 I38153	gene retin protein
35	93.5	4.6	791	2 A46140	diacylglycerol kin
36	93.5	4.6	796	2 H46140	diacylglycerol kin
37	92.5	4.6	775	2 H84567	probable diacygly
38	92.5	4.6	732	2 T16422	hypothetical prote
39	92	4.6	782	2 T16422	hypothetical prote
40	92	4.6	1192	2 T17255	semaphorin C - mou
41	91.5	4.5	304	2 A86842	hypothetical prote
42	91.5	4.5	364	2 H87212	conserved hypotet
43	91.5	4.5	1172	2 T36053	probable glucose e
44	91	4.5	219	2 AE0724	probable ABC-type
45	90.5	4.5	490	2 G85354	probable hydrolase
					hypothetical prote

ALIGNMENTS

RESULT 1
T05162
hypothetical protein F18E5.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05162
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dijkse, W.; Stiekema, W.; Bancroft, I.;
submitted to the Protein Sequence Database, August 1998
A:Reference number: 215400
A:Accession: T05162
A:Molecule type: DNA
A:Residues: 1-1240 <REV>
A:Cross-references: EMBL:AL022603
A:Experimental source: cultivar Columbia; BAC clone F18E5
C:Genetics:
A:Map position: 4
A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1
A:Note: F18E5.160

Query Match	22.4%	Score 451;	DB 2;	Length 1240;
Best Local Similarity	31.9%	Pred. No. 1e-31;		
Matches: 122;	Conservative 63;	Mismatches 127;	Indels 70;	Gaps 13;
QY 11	LPRPRVYLILNPRGKGRKALQFRSHVQPLAEIDISTMLTERRHARLVSEELG 70			
DB 374	LGRPRLLVFNVPFGSKRSAREIFVKEVKPLFEDADVLEIQETKYQLHAKKFKSMDS 433			
QY 71	RMDALVMSGDGLMEHVNGLMERPDWETAIQKPLCSLPAGSGNALAASLNHYAGVEQVT 130			
DB 434	KYDGIVCVSGDGLIEVYVNGLLERADWRNAKRLPGMPAGNGMIMSLDTVGLRCA 493			
QY 131	NEDLLTNCITLLCRRLSLSPNNLSLHTASGLRSLFSLAMGFADVDLESEKRYRLGEM 190			
DB 494	N-----SMTISIRGHRKRSVDVATL-AOCNTEFFSVMLAMGLADIDESEKPFMWGSA 547			
QY 191	RTTLGTFL-----RLAALRTYNGRLALYLV-GRVSKTASAVVVO- 230			
DB 548	RIDFVYVCLVDKFDNVCIAVVKLALQRTICLRNRYGRILFLPAPFEGGQAPSSLVQE 607			
QY 231	-----OGP-----VDAILVPLLEEPVSHMTVVPDEDFVLVIALHS-HLGSE-MFA 274			
DB 608	PHVSKVEYVGQPFKEFKFEDLEKREMKGFPVTW-----LHNVMGSENTLT 653			
QY 275	APMGSCAGVMHLEFVRAGVSRAMLRLFLANEGKRHMEYECPLYLVY-----PVAFRL 329			
DB 634	APAAKFSQGYLDLIVLK-NCPKLVLISLMRQSSGTHV--ESPVIYVIKLVEKKAFAVL 710			
QY 330	EP-----KDKGVFAVDCGLM 345			
DB 711	EPGALVDEPKDEGIIDSDGEVL 732			

Query Match 7.7%: Score 155.5; DB 2: Length 306;
 Best Local Similarity 22.2%; Pred. No. 4.2e-06;
 Matches 77; Conservative 66; Mismatches 129; Indels 75; Gaps 16;

OY 16 RYVLLNPGCGKALQFRSHVOPLLAEISFTMLTERNNHARELV--RSEELGRMD 73
 Db 4 KAMITNPAAGNKKRKL--PDAEKILTNMDEVTLPSPAPKSTTLAQAQAEAG-TE 61
 OY 74 ALVYMGDDGIMHEVYNGME--RPDWEFAIQPCSLPAGSGNALASLNHYAGYEQVT 130
 Db 62 VTIAGGDDGTVNEVNGMLQVRRP-----KLGILPVGTTNDYARALN----- 105
 OY 131 NEDLLTNCITLLCRLLSPMNLISLHTASGLRLFSVLSLAWGFADVDLESEKYRLGEM 190
 Db 106 AKQPLEALQIIRKQETIR---VDIGKANETFEFF--INNAAGRIITEI----- 147
 OY 191 RFTLGTFLRLAALRTYRGRRLAYLPVGRVSGKTPASPVVVOGQPVDAHLVPLEEPVPSHMT 250
 Db 148 -----TYAVKESKMSKMGRLAYLFSGLT-----VLEKRLSPVNVETIYNDE----- 187
 OY 251 VPDDEDFVLALLSHLHSGSEMFAPMGRCAAGVNHLYFRAGVSRAMLRLFLAMEKGR 310
 Db 188 IFKGEILFEFVNKNTSVGSMETLCP--AQLNSGMPELLIKK--VSPKRLQFLFASTIKGT 245
 OY 311 HMEYECRYLVYVVAFAFLPEKDKGVFAVDGELMVS--EAYOGQVHP 356
 Db 246 HL--NSPHVIAH-----RTNKVTIKSDADLNTSYDQVYCGKAP 281

RESULT 8

Conserved hypothetical protein lmo2557 [imported] - *Listeria monocytogenes* (strain EGD-e AE1394)
 C:Species: *Listeria monocytogenes*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C/Accession: AE1394
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karet, U.
 Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kunz, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Trierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AE1394
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-306 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAP00635.1; PID:g16412045; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics: lmo2557

Query Match 7.0%: Score 140.5; DB 2: Length 306;
 Best Local Similarity 21.7%; Pred. No. 9.2e-05;
 Matches 76; Conservative 60; Mismatches 132; Indels 83; Gaps 16;

OY 16 RYVLLNPGCGKALQFRSHVOPLLAEISFTMLTERNNHARELV--RSEELGRMD 73
 Db 4 KAMITNPAAGNKKRKL--PDAEKILTNMDEVTLPSPAPKSTTLAQAQAEAG-TE 61
 OY 74 ALVYMGDDGIMHEVYNGME--RPDWEFAIQPCSLPAGSGNALASLNHYAGYEQVT 130
 Db 62 VTIAGGDDGTVNEVNGMLQVRRP-----KLGILPVGTTNDYARALN-----FARDPL 110
 OY 132 EDLLTNCITLLCRLLSPMNLISLHTASGLRLFSVLSLAWGFADVDLESEKYRLGEM 190
 Db 111 EAL-----RIIAQETIRVDICKANETEFINNAGRIITEI----- 147
 OY 191 RFTLGTFLRLAALRTYRGRRLAYLPVGRVSGKTPASPVVVOGQPVDAHLVPLEEPVPSHMT 250
 Db 148 -----TYAVKESKMSKMGRLAYLFSG--LTVLPLSPVVE-----I 182

OY 251 VPDDEDFVLALLSHLHSGSEMFAPMGRCAAGVNHLYFRAGVSRAMLRLFLAMEKGR 306
 Db 183 AYNDLFEFKILLFEVKNSSVGMETLCPAPALNSGMPELLIKK--VSPKRLQFLFASTI 241
 OY 307 EGRHMEYECRYLVYVVAFAFLPEKDKGVFAVDGELMVS--EAYOGQVHP 356
 Db 242 KQVTHL--NSPHVIAH-----RTNKVTINSQADLNTSYDQVYCGKAP 281

RESULT 9

Conserved hypothetical protein ytlir - *Bacillus subtilis* H69995
 C:Species: *Bacillus subtilis*
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C/Accession: H69995
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Enllich, S.D.; Emerson, P.T.; Entian, K.D.; Erlington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mei Reger, M.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portere Yager, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl A:Authors: Schleib, S.; Schreier, R.; Scoffone, F.; Seguchi, J.; Sekosaka, A.; Se akeuchi, M.; Tomakoshi, A.; Tanaka, T.; Terpestre, P.; Tognoni, A.; Tosato, V.; Uchida T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamamoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Denchik, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis* A:Reference number: A69360; MUID:98044033; PMID:9384377
 A:Accession: H69995
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-309 <KUN>
 A:Cross-references: GB:699119; GB:AL009126; NID:92635411; PIDN:CA14972.1; PID:e11858
 A:Experimental source: strain 168
 C:Genetics: ytlir

Query Match 6.8%: Score 136.5; DB 2: Length 309;
 Best Local Similarity 23.4%; Pred. No. 0.00021;
 Matches 79; Conservative 47; Mismatches 148; Indels 63; Gaps 17;

OY 20 LNPNGCGKALQFRSHVOPLLAEISFTMLTERNNHARELVRS---EELGMDAL 75
 Db 7 11NPTAGHRNGRLRWKS--IQELIKRKEVRSFLETHPHCAEVLARQISTIOGY-RLKRL 64
 OY 76 VYMGDDGIMHEVYNGME--RPDWEFAIQPCSLPAGSGNALASLNHYAGYEQVTNEDLL 135
 Db 65 IYIGDDGIMHEVYNGMLQVDDIE-----LSFVPAGAYNDSPRFS-----IKKIDLI 111
 OY 136 TNCITLLCRLLSPMNLISLHTASGLRLFSVLSLAWGFADVDLESEKYRLGEMFTLG 195
 Db 112 QE-----IKKKRRLT--RTFLGS--VNFLODKSOLITFMNIGIGCPAYVKKAMAFPLR 164
 OY 196 -----TLRLAALRTYRGRRLAYLPVGRVSGKTPASPVVVOGQPVDAHLVPLEEPVPSH--WT 250
 Db 165 RYVLLRL-----RFLVYPLSHLASATKRPFLACTED-----ETREPDVWF 209
 OY 251 VPDDEDFVLALLSHLHSGSEMFAPMGRCAAGVNHLYFRAGVSRAMLRLFL--AMEK 308
 Db 210 AVVSN-----HPPYGGGMAAPLAPNPKRTDVIYF--NPFILKRYVLLCLMAF 257
 OY 309 GRHMEYECRYLVYVVAFAFLPEKDKGVFAVDGELM 345
 Db 258 GKRTMDGVTFKAKDITF--YKRD--KIPFIADGEIM 291

RESULT 10

Conserved hypothetical protein BH1953 [imported] - *Bacillus halodurans* (strain C-125) A83894
 C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: A83894
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A83894
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05672.1; GSPDB:GN0C
C:Genetics:
A:Gene: BH1953

Query Match 6.4%; Score 130; DB 2; Length 295;
Best Local Similarity 22.0%; Pred. No. 0.00076;
Matches 78; Conservative 50; Mismatches 138; Indels 88; Gaps 14;

QY 19 VLINPRGKRALQLFRRHVOPPLAEAFISPTLMTERRNHARELVRSSEELGRMDALVVM 78
D 4 LIIKASGNGKGGRTWK-VEEIQIRNTPYLVFTSGSHATITVKELLEGVYTIIV 62
QY 79 SGDLMEHVNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGTEQYTNEDLTNC 138
D 63 GGDGTINEVANGLVNH-----RVPGLIIPAGSGNDFARCLNTPMHEKALHR----- 109
QY 139 TLLCRRLSPMLSLHTASGLRFSVLSLAWCFIADV--DLESEKRR-----LGE 190
D 110 ---LFEKKOKKVDLHLGQRCLVTGT-----GPDGIATVTEATIKMNFNGFGL 161
QY 191 RFTLGTFLRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDVAHLVPLEEPVSHMT 250
D 162 SYVLS---MLEVLDKDY-----PTNIQIT 182
QY 251 VVPDEDF---VLVIALHLS-HLGEEMFAAPKRCAGVHLEFYRAGVSRMLRLFLIAM 306
D 183 VDGELEFSGVLAIVANSPVGGIRICPEASVDDGLNICVNA-GMSKQMLRLFPKA 241
QY 307 EKGRR--MEYECPLY---VVVPAVAFRLPEKDGFAVAGELVSSAVOQOV 354
D 242 YKGRHVMEQHTLTGKDYV-----OSDIPVLVOSDEPIESPRLQI 287

RESULT 11
F69795
conserved hypothetical protein yero - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C/Accession: F69795
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivoita, C.; Roche, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, J.; Tosi, V.; Uchiyama,
T.; Winkler, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:96044033; PMID:9384377
A:Accession: F69795
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-303 <KUN>
A:Cross-references: GB:299107; GB:AL009126; NID:g2632866; PIDN:CAB12492.1; PID:ell182652;
A:Experimental source: strain 168
C:Genetics:

A:Gene: yero

Query Match 6.4%; Score 129.5; DB 2; Length 303;
Best Local Similarity 20.6%; Pred. No. 0.00087;
Matches 68; Conservative 44; Mismatches 113; Indels 105; Gaps 13;

QY 16 RVVLINPRGKRALQLFRRHVOPPL---AEAFISPTLMTERRNHARELVRSSEELGRW 72
D 3 RARILYNPTSGR---EIFKKHIAQVIOKFERAGYETSHATTCAGDTAAKFAALREF 58
QY 73 DALVMSDDGLMEHVNGLM---ERPDWETAIOKPLCSLPAGSGNALAASLNHYAGTEQY 129
D 59 DLIAGGDTINEVANGLVNH-----RVPGLIIPAGSGNDFARCLNTPMHEKALHR----- 103
QY 130 TNEDLTNCITLLCRRLSPML-----LSLHTASGLRFSVLSLAWCFIADVPLESEK 183
D 104 PREDILKADIVY-NGVARPIDIGVNGQYFNTNAGGRLTEL-----TYVPSKL 153
QY 184 YRLGEMFTLQFLRLAALR-----TYGRRLAYLPVGRVSKTPASPVVVOGPDVAH 237
D 154 KTMGLQVAYLKGMELPSLRTEVEIEYDKLF-----QGEIMLF 194
QY 238 LVPLEEPVSHMTVVPDEDF-----VLVIALHLSHGEMFAAPM 277
D 195 LVTLNVSQGEFKLAPDSSLDGMDLMIKKANLAETIRATATMLRGEHINDQ----- 248
QY 278 GRCAGVNHLFY-----VRAGVSRMLRL 302
D 249 -----HIITKANRYKANVSEKQMLNL 270

RESULT 12
D86677
transcription regulator [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C/Species: *Lactococcus lactis* subsp. *lactis*
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86677
R:Bojoltin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh
genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: D86677
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <STO>
A:Cross-references: GB:AE005176; PID:g12723295; PIDN:AAK04518.1; GSPDB:GN0146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yecE

Query Match 6.3%; Score 127; DB 2; Length 342;
Best Local Similarity 20.4%; Pred. No. 0.0017;
Matches 72; Conservative 55; Mismatches 122; Indels 104; Gaps 15;

QY 19 VLINPRGKRALQLFRRHVOPPLAEE---EISFTLMTERRNHARELVRSSEELGRWD 73
D 6 LYNPTSG---EIKKYIADILDKLEGYEAASVYQTAQDSAKKRAARATEAG-FD 60
QY 74 ALVMSGDLMEHVNGLM---ERPDWETAIOKPLCSLPAGSGNALAASLNHYAGTEQY 130
D 61 LIIAGGDTINEVANGVSPFEKRPD-----LAIYPTGTTNDFARLAKIPRG----- 107
QY 131 NEDDLTNCITLLCRRLSPMLSLHTASGLRFSVLSLAWCFIADVPLESEKRYRLGEM 190
D 108 -----KPLEAIEI-----IKKNOITLNDVGHAVITRFQDDQY----- 139
QY 191 RFTLGTFLRLAALRTYGRRLAYLPVGRVSKTPASPVVVOGPDVA 236
D 140 -----FNTIAGGGTTELTVSVPSHLKTAFLYLAAG-----A 174
QY 237 HLVPLEEPVSHMTVVPDE---DFVLVIALHLSHGEMFAAPMGRCAGVNHLYFR 291
D 175 ELLPVKRAVPR--VVHDEGVPEGDISMFEFALTNVSQGEFKIADPAKLDLGLFTLIVK 232

OY 292 AGVSRAMLRLFLAMEKGRHMEYECPLYVYVAVAFRLKPKDKGVFA-VDGE 343
 Db 233 TD-LFELIALALAVANGKHL--DDVNLLEYIKTSKIELEALOGKILINDGE 282

RESULT 13

Transcription regulator [Imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: F86849
 R:Bioclin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.s.
 A:Reference number: AB6625; MUID:21235186; PMID:11337471
 A:Accession: F86849
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <STO>
 A:Cross-references: GB:AE005176; PID:g12724823; PIDN:AAK05896.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:

Query Match 6.2%; Score 126; DB 2; Length 302;
 Best Local Similarity 21.2%; Pred. No. 0.0018;
 Matches 75; Conservative 64; Mismatches 123; Indels 92; Gaps 17;

OY 19 VLLNPRGGKALQLEFRSHVOPFLAEISFTLMLTERNHARELYRS--ELGKMDALV 76
 Db 5 LLAIPNSGAKGKATLETL--PLEKNNYERLEKTAAGECALRQILDKNPDHVL 63
 OY 77 VMGGDLMEVYVNGLMERPDMETAIQKPLCSLPAGSMNLAASL-----NHVAGYEQ 128
 Db 64 ILGGDGLSLVINELEPE-----EENAFSTIPSGSNDFAHSKLIKALDPLESEPARRG 115
 OY 129 VTNEDLITNCTLLCRLLSPNLSLHTASGLRSLVSLAMGFADVDSEKRYRLG 168
 Db 116 INHIFITM-----YQSGLSGIVLNIGIGDALT--VKSANEKRLK 156
 OY 189 EM--RFLTGFLMALRYRGRALYLPVGVGSKTPASVYVVOGPDAILVPLEEVP 246
 Db 157 QVLMKLKLSF-----SYL-----LTALHVLITKRPFA-LIEVEN--- 191
 OY 247 SHMTVYDEDFVYLALLSHLSGSEMPFAWRCAQVMHLEFYVAGVSRAMLRLFL-- 304
 Db 192 ---QEISLENAFLMTTKTPYFGGVKISPEATMENADITL---VEYKHHLLATSLI 244
 OY 305 -AMKGRHMEYECPLYVYVAVAFRLKPKDKGVFAVDGEIYVAVO--GGVH 355
 Db 245 PSYLRKHMLH-----PLFLHVRVSPQ-----FSV--ELALSQVVDIGDGIH 283

RESULT 14

hypothetical protein homolog lln1865 [Imported] - Listeria innocua (strain C11p11262)
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC1665
 R:Gibson, F.; Frenkel, I.; Buchrieser, C.; Amend, A.; Bagnato, F.; Berche, P.; Bloeker
 K.; Glaser, F.; Hengstler, G.; Duchaud, E.; Durand, L.; Dussutelle, O.; Entian, K.D.; Fsihi, H.
 J.; Dominguez Jernall, G.; Duchaud, E.; Durand, L.; Dussutelle, O.; Entian, K.D.; Fsihi, H.
 Science 294, 849-852, 2001
 A:Authors: refc, J.; Kuhn, M.; Kunst, F.; Kurpkat, G.; Madueno, E.; Maitounnam, A.; Me
 O.; Schuster, T.; Sines, N.; Tlorrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of Listeria species
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1665
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-310 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97095.1; PID:g16414366; GSPDB:GN00178

A:Experimental source: strain C11p11262

C:Genetics:

A:Gene: lln1865

Query Match 6.2%; Score 124.5; DB 2; Length 310;
 Best Local Similarity 21.3%; Pred. No. 0.0025;
 Matches 74; Conservative 55; Mismatches 118; Indels 101; Gaps 15;

OY 19 VLLNPRGGKALQLEFRSHVOPFLAEISFTLMLTERNHARELYRS--ELGKMDALV 70
 Db 7 VLYNTSR-----ELKKNLADVLSTLEBQAGVTSALHATPEPDAAHAAEEAVNR-- 58
 OY 71 RMDALVWVSGDLMEVYVNGLMERPDMETAIQKPLCSLPAGSMNLAASLNHVAGYEQT 130
 Db 59 RYDVVAAGGDTINEVINCIAQP-----YRPKGIIPGTCTNDF----- 100
 OY 131 NEDLITNCTLLCRLLSPNLSLHTASGLRSLVSLAMGFADVDSEKRY----- 184
 Db 101 -----HALHVRDV-----IKATKILIAQSVAMDICRANTDTPFINIG 139
 OY 185 -RELGEHFTLCTFLRLAALRYRGRALY-----LPVGRVSGSKTPASVYVVOGPDVA 236
 Db 140 GGRLEFLVDP-----SRKTMIGQLAYTLKGIEMF-----SLKATKVVVEYDQ-- 186
 OY 237 HLVPLEDFEVPVSHMTVYDEDFVYLALLSHLSGSEMPFAWRCAQVMHLEFYVAGVSR 296
 Db 187 -----VFGEVWFFLGLTNSIGGFKEI--APDAKLDGCKFSLIVKK--VM 230
 OY 297 AMHLRLFLAMEKGRHMEYECPLYVYVAVAFRLKPKDKGVFAVDGEI 344
 Db 231 AEFILVTLALRGDIKE--PNVIYKSEKSVSHSD--KMLINDGEL 275

RESULT 15

multidrug resistance protein bmr1 [Imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: F83871
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: F83871
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-311 <STO>
 A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05493.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:

Query Match 6.1%; Score 122.5; DB 2; Length 311;
 Best Local Similarity 30.5%; Pred. No. 0.0038;
 Matches 32; Conservative 17; Mismatches 47; Indels 9; Gaps 3;

OY 16 RVLLNPRGGKALQLEFRSHVOPFLAEISFTLMLTERNHARELYSEELGRDAL 75
 Db 4 RAVLLNCKAGANGKT--DKYLDKDVFRALQSIGELTTPSEKGDSEISCEEREV---DL 59
 OY 76 VVWVSGDLMEVYVNGLMERPDMETAIQKPLCSLPAGSMNLAASL 120
 Db 60 IVMGGDTVHECTINGLAVSPS-----PPPLVLPFGTCNDFAFSL 59

Search completed: May 9, 2003, 16:57:23
 Job time : 26 secs

JOURNAL MEDLINE 20263733
 PUBMED 10802064
 REFERENCE 2 (bases 1 to 1783)
 AUTHORS Nava, V. E., Lacana, E., Poulton, S., Liu, H., Suglione, M., Kono, K.,
 TITLE Miltien, S., Kohama, T., and Spiegel, S.
 JOURNAL Submitted (23 FEB-2000) Biochemistry, Georgetown University, 3900
 Reservoir Rd NW BSB km 357, Washington, DC 20007, USA
 FEATURES Location/Qualifiers
 source 1..1783
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q25.2"
 337..1491
 /codon_start=1
 /product="sphingosine kinase-1"
 /protein_id="AF73423.1"
 /db_xref="GI:8132868"
 /translation="MDPAGCPRGVLPKPCRVLLNDRGKGKALQFESHVQPLAE
 AEISFTMLTERNHARELYSEELRWDALVMSGDLNHEVYNGIMRPMETAIO
 KPLCSLPGAGNALASLNHYAGEOVNEDLLNCTLLCRBLSPMLLSLHNSG
 LRLSPSLAMGFIADVDLESEKRYRLEMEFTLTGLRLALRTYRGSLAYLPGAR
 GSKPASPVPVVOGDAHVPLEEPVPSHMTWNPDEPFLIALHSHLSGEMAAP
 MRCAGVMLEFYRAGVSRAMLLRPLAKRGMHEPDPYLVYPPVAVPLEEKD
 KCFVAVDCELMSEYAVOGVHNPYFMVSGVBPSPMKRQMPPEBRL"
 BASE COUNT 314 a 546 c 591 g 332 t
 ORIGIN
 Query Match 98.6%; Score 1550.8; DB 9; Length 1783;
 Best Local Similarity 99.5%; Pred. No. 7; Le-2/79;
 Matches 1366; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 1 GCCCCACAGCCGCGCTTCGACGCGCCCTGGGACACCCGATTAAGAGCTGAAGGACG 60
 207 GCCCCACAGCCGCGCTTCGACGCGCCCTGGGACACCCGATTAAGAGCTGAAGGACG 266
 61 AGCCGCCGCCAGGCGGACGCG-CCCCACAGCGCCAGGAGACCCCTGGCAGCGGAGCGCG 119
 267 AGCCGCCGCCAGGCGGACGCGCCGCCACAGCGCCAGGAGACCCCTGGCAGCGGAGCGCG 326
 120 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 179
 327 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 386
 180 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 239
 387 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 446
 240 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 239
 447 GGTGAGGTTATGATCCAGCGGCGGCGCCCGGCGGTGTCTCCCGCGCCCTGCGCGT 506
 300 GAACACCGCGCGGACGCTGTGCGGTGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 359
 507 GAACACCGCGCGGACGCTGTGCGGTGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 566
 567 GAACACCGCGCGGACGCTGTGCGGTGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 626
 627 GAACACCGCGCGGACGCTGTGCGGTGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 686
 480 AGCTTCCTTGAACCATTAATGCTGCTATGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 539
 687 AGCTTCCTTGAACCATTAATGCTGCTATGAGAGAGCTGGCGCGCTGGAGCGCTGTGCGT 746
 540 CTGGAGGCTATTGCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
 747 CTGGAGGCTATTGCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806

600 GCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 807 GCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 660 GCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 867 GCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 720 CCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 927 CCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 780 AGTGGGTTCCAAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
 987 AGTGGGTTCCAAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 840 CCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
 1047 CCGTTCGCGGCGCGCGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 900 TGTGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
 1107 TGTGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
 960 GGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 1167 GGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
 1020 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
 1227 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1286
 1080 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
 1287 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346
 1140 GTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
 1347 GTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406
 1200 CTACTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 1407 CTACTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
 1260 GCGACGCGCAAG 1319
 1467 GCGACGCGCAAG 1526
 1320 AGGACG 1379
 1527 AGGACG 1586
 1380 GAGACTTCCTGCGAG 1439
 1587 GAGACTTCCTGCGAG 1646
 1440 TGAAGTCTGCGTCAAG 1499
 1647 TGAAGTCTGCGTCAAG 1706
 1500 TTTTGTTCGAG 1559
 1707 TTTTGTTCGAG 1766
 1560 AAAAAAAAAAAAAA 1573
 1767 AAAAAAAAAAAAAA 1780

RESULT 2
 BC008040 1824 bp mRNA linear PRI 12-JUL-2001
 LOCUS BC008040

DEFINITION Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
ACCESSION BC008040
VERSION BC008040.1 GI:14165485
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1824)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahay, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IMAGE Plate: 24 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10433790.
FEATURES
source
1..1824
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15041 IMAGE:3831657"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH-MGC-20"
/lab_host="DH10B-R"
/note="vector: pOTB7"
339..1493
/codon_start=1
/product="Unknown (protein for MGC:15041)"
/protein_id="AAH08040.1"
/db_xref="gi:14165485"
/translation="MDPAGGRGVLPKRCRLVLLNPRGGKKAIOIFRSHVPLAE
AEISFTLMTERNRLARLVREELGRVDALVMSGDGIMHEVVGIMERPDMETAIO
KPLCSLPAGSSNALASLNHYAGYEDVETNEDLITCTLLICRLLSPNNLLSLHTASG
LREFVLSLWGFIDVDLESKRYRLGEMRTLTGTFLLAALRTYRGRLAYLPVGRV
GSKTPASPVVVOGPDVAHIVPLEBPVPSHMTVVPDDEPVLIALIHSIGSEMPAAR
MGRCAAGYMHLYFYRAGYSRANLELFLAMEGRIMEVECPILVYVYVAFLELRKDG
KGVRVNDDELINSEAVVGVQVHPNFTWMSGVCEVPPSPKPMQMPPEEPL"
CDS
BASE COUNT 354 a 551 c 592 g 327 t
ORIGIN
Query Match 98.5% Score 1549.2; DB 9; Length 1824;
Blast Local Similarity 99.4% Pred. No. 1.4e-278;
Matches 165; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GCCCCACACGCGCCCTGCGACGCCCTGGGACAGCAGTAAGAGCTGAAGCAGG 60
DB 209 GCCCCACACGCGCCCTGCGACGCCCTGGGACAGCAGTAAGAGCTGAAGCAGG 268
QY 61 AACCAGCGCCACGCGGAGCG-CCCCACAGCGCCAGGAGCCCTCTGACGCGGAGCGCG 119
DB 269 AACCAGCGCCACGCGGAGCGCGCCACAGCGCCAGGAGCCCTCTGACGCGGAGCGCG 328
QY 120 GGTGCGAGTTATGATGATCAAGCGGCGCGCCCGGCGGCTGCTCCCGCGCGCTGCGCG 179
DB 329 GGTGCGAGTTATGATGATCAAGCGGCGCGCCCGGCGGCTGCTCCCGCGCGCTGCGCG 388
QY 180 GGTGCTGTCTGTAACCCCGCGCGCGCAAGGCAAGGCTTGCACCTCTCCGAGAGTCA 239

DB 389 CCTGTGCTGTGTAACCCCGCGCGCGCAAGGCAAGGCTTGCACCTCTCCGAGTCA 448
QY 240 CGTCGAGCCCGCTTTGGCTGAGAGCTGAANTCTTCACAGCTGATGCTACGAGGCGG 299
DB 449 CGTCGAGCCCGCTTTGGCTGAGAGCTGAANTCTTCACAGCTGATGCTACGAGGCGG 508
QY 300 GAACACGCGCGGAGAGCTGTGCGGTCGAGAGAGCTGCGCCCTGGAGAGCTGTGTGT 359
DB 509 GAACACGCGCGGAGAGCTGTGCGGTCGAGAGAGCTGCGCCCTGGAGAGCTGTGTGT 568
QY 360 CATGCTGTGAGAGAGCGGCTGATGACAGAGGTGTGAACGGGCTCATAGAGCGGCTGACTG 419
DB 569 CATGCTGTGAGAGAGCGGCTGATGACAGAGGTGTGAACGGGCTCATAGAGCGGCTGACTG 628
QY 420 GAGAGCCGCTATCCAGAACACCCCTGTGTAGCTCCAGAGAGCTGTGGAACGCGCTGCG 479
DB 629 GAGAGCCGCTATCCAGAACACCCCTGTGTAGCTCCAGAGAGCTGTGGAACGCGCTGCG 688
QY 480 AGCTTCCTTGAACCATTTATGCTGAGTATGAGAGAGGTCACCAATGAAAGACCTCTGACCA 539
DB 689 AGCTTCCTTGAACCATTTATGCTGAGTATGAGAGAGGTCACCAATGAAAGACCTCTGACCA 748
QY 540 CTGCACGCTATTGCTGTGCGCGCGGCTGCTGCACCCATGAACCTGCTGCTGTCACAC 599
DB 749 CTGCACGCTATTGCTGTGCGCGCGGCTGCTGCACCCATGAACCTGCTGCTGTCACAC 808
QY 600 GCGTTGCGGCGCTGCGGCTCTTCTGTGCTGACAGCTGCGGCTTATGCTGATGT 659
DB 809 GCGTTGCGGCGCTGCGGCTCTTCTGTGCTGACAGCTGCGGCTTATGCTGATGT 868
QY 660 GACACTAGAGAGTGAAGATATGAGGCTGCGGAGAGATGAGGCTTACCTGAGGACCTT 719
DB 869 GACACTAGAGAGTGAAGATATGAGGCTGCGGAGAGATGAGGCTTACCTGAGGACCTT 928
QY 720 CCTGCTGTGCGACGCCCTGCGACCTACCGCGGCGGCTGACCTGCTGCTGAGAG 779
DB 929 CCTGCTGTGCGACGCCCTGCGACCTACCGCGGCGGCTGACCTGCTGCTGAGAG 988
QY 780 AGTGGGTTCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
DB 989 AGTGGGTTCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
QY 840 CCTGTGCGACAGTGAAGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
DB 1049 CCTGTGCGACAGTGAAGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108
QY 900 TGTGCTAGTCTGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
DB 1109 TGTGCTAGTCTGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1168
QY 960 GCGCGCTGTCGACGCTGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
DB 1169 GCGCGCTGTCGACGCTGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1228
QY 1020 CATGCTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
DB 1229 CATGCTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1288
QY 1080 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
DB 1289 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
QY 1140 GTTTCAGTGTGATGAGGAAATGATGTTAGCGAGGCTGTGACAGGCTGAGTGCACCCAAA 1199
DB 1349 GTTTCAGTGTGATGAGGAAATGATGTTAGCGAGGCTGTGACAGGCTGAGTGCACCCAAA 1408
QY 1200 CTACTTGTGATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
DB 1409 CTACTTGTGATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1468
QY 1260 GCCACCGCCAGAGAGAGCCCTTATGACCCCTGCGGCGGCTGCTGCTTACTTACTTGC 1319

Db 1469 GCCACCCGCAAGAGCCCTTATGACCCCTGGGCGCCGCTGTGCTTACTTCTTC 1528

Oy 1320 AGGACCCCTTCCTCTTCCCTAGGCGTGTGACAGGCTGTCTGTGGGGTGGAG 1379

Db 1329 AGGACCCCTTCCTCTTCCCTAGGCGTGTGACAGGCTGTCTGTGGGGTGGAG 1588

Oy 1380 GAGACCTCTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1439

Db 1389 GAGACCTCTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1648

Oy 1440 TGAAGCTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1499

Db 1649 TGAAGCTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1708

Oy 1500 TTTGTTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1559

Db 1709 TTTGTTCTGAGAGGCTGAGAGGCTGAGAGGCTGTGGGGGAGACAGCCAGAA 1768

Oy 1560 AAAAAAAAAAAAAA 1573

Db 1769 AAAAAAAAAAAAAA 1782

RESULT 3
AK022402
LOCUS
DEFINITION
AK022402 1821 bp mRNA linear PRI 01-AUG-2002
Homo sapiens CDNA FLJ12340 fls, clone MAMMA1002268, moderately
similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
AK022402
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Mishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
Wagatsuma, M., Hosokawa, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
Oho, Y., Takiguchi, S., Matsubara, S., Kimura, K., Murakami, K.,
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Nakamatsu, A.,
Nakamura, Y., Nagahata, K., Masuno, Y., Ninomiya, K., and Iwayanagi, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1821)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
SOURCE
1. 1821
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1002268"
/cisseq_type="Mammalian gland"
/clone_id="MAMMA1"
/note="Cloning vector: pME18SFL3"
393..1547
/note="unlabeled protein product"
/codon_start=1
/protein_id="BAB14028.1"
/db_xref="GI:10433791"
/translation="MDPAGGPRGVLPKPRVLYLNLPRGGKALQLFRSHVPLLAIE

BASE COUNT 301 a 569 c 614 g 337 t

ORIGIN

Query Match 98.3%; Score 1547; DB 9; Length 1821;
Best Local Similarity 99.9%; Pred. No. 3.6e-278;
Matches 1558; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 GCCCCACAGCCGCGCTGACAGCCGCTGGGCGAGCAGCAGTAAAGAGCTGAAGCAGG 60
|||||
263 GCCCCACAGCCGCGCTGACAGCCGCTGGGCGAGCAGCAGTAAAGAGCTGAAGCAGG 322
|||||
61 AGCCGCGCGCAGCGGCGAGCG-CCGACAGCGCGCAGGAGCCGCTGCGAGCGGAGCGCG 119
|||||
323 AGCCGCGCGCAGCGGCGAGCGCCGCGACAGCGCGCAGGAGCCGCTGCGAGCGGAGCGCG 382
|||||
120 GCTGCGAGCTTATGATCCAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
|||||
383 GCTGCGAGCTTATGATCCAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 442
|||||
180 GCTGCTGCTGCTGAACCCG 239
|||||
443 GCTGCTGCTGCTGAACCCG 502
|||||
240 GCTGCGAGCCGCTTTGGCTGAGGCTGAATCTCTTACGCGTATGCTGCTGAGCGCGCG 299
|||||
503 GCTGCGAGCCGCTTTGGCTGAGGCTGAATCTCTTACGCGTATGCTGCTGAGCGCGCG 562
|||||
300 GAACCAAGCG 359
|||||
563 GAACCAAGCG 622
|||||
360 CAGTCTGAGAGCG 419
|||||
623 CAGTCTGAGAGCG 479
|||||
420 GAGAGCG 529
|||||
683 GAGAGCG 582
|||||
480 AGCTGCTTGAACCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
|||||
743 AGCTGCTTGAACCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
|||||
540 CTGCGAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
|||||
803 CTGCGAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
|||||
600 GCTGCGAGCG 819
|||||
863 GCTGCGAGCG 922
|||||
660 GAGAGCTTGAACCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
|||||
923 GAGAGCTTGAACCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
|||||
720 CTTGCTGCTGCGAGCG 779
|||||
983 CTTGCTGCTGCGAGCG 1042
|||||
780 AGTGGGTTCAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
|||||
1043 AGTGGGTTCAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102
|||||
840 CTTGCTGCTGCGAGCG 899
|||||
1103 CTTGCTGCTGCGAGCG 1162
|||||

QY	791	AGACACCTGACCTCCCGGTTGTGGTCAGACGAGGCGCGGTGATGACACACCTTGTGGCAC	850
Db	931	AGACACCTGACCTCCCGGTTGTGGTCAGACGAGGCGCGGTGATGACACACCTTGTGGCAC	990
QY	851	TGAGAGGACCAAGTGCCTCTCTACTGTGACAGAGTGTGTGCCGACGAGAGACTTTGTGCTAGTCC	910
Db	991	TGAGAGGACCAAGTGCCTCTCTACTGTGACAGAGTGTGTGCCGACGAGAGACTTTGTGCTAGTCC	1050
QY	911	TGCGACTGCTCACTCGCACCCTGGGCAAGTGAATGTTTGTCTGACACCCATGGGCGCTGTG	970
Db	1051	TGCGACTGCTCACTCGCACCCTGGGCAAGTGAATGTTTGTCTGACACCCATGGGCGCTGTG	1110
QY	971	CAGGTGGGTCATGATCTGTCTTACAGTGGGGGGGGGAGTGTCTCTGCTCCATGCTGTGCTG	1030
Db	1111	CAGGTGGGTCATGATCTGTCTTACAGTGGGGGGGGGAGTGTCTCTGCTCCATGCTGTGCTG	1170
QY	1031	GCCCTTCTCTGCGCATGAGAGAGGCAAGGCATATGAGATATGAAATGCCCTACTTGGTAT	1090
Db	1171	GCCCTTCTCTGCGCATGAGAGAGGCAAGGCATATGAGATATGAAATGCCCTACTTGGTAT	1230
QY	1091	ATGTGCCCCGTGGTGCCTTCCGCTTTGGAGCCCCANAGATGGAAAGTGTGTTTGCAGTGG	1150
Db	1231	ATGTGCCCCGTGGTGCCTTCCGCTTTGGAGCCCCANAGATGGAAAGTGTGTTTGCAGTGG	1290
QY	1151	ATGGGGAAATTCATGTTATGAGAGAGGCGCGGAGAGGCGAGAGTGCACCCAACTACTTCTGGA	1210
Db	1291	ATGGGGAAATTCATGTTATGAGAGAGGCGCGGAGAGGCGAGAGTGCACCCAACTACTTCTGGA	1350
QY	1211	TGCTCAGCGGTTGGGTGGAGGCCCGCGCCACACTGGAAGGCCCCAGACAGATGCTCACCGCCAG	1270
Db	1351	TGCTCAGCGGTTGGGTGGAGGCCCGCGCCACACTGGAAGGCCCCAGACAGATGCTCACCGCCAG	1410
QY	1271	AAGAGCCCTTATGACACCCCTGGGCGCGCTGTGCTTAACTGCTACTTGCAGAGACCTTCC	1330
Db	1411	AAGAGCCCTTATGACACCCCTGGGCGCGCTGTGCTTAACTGCTACTTGCAGAGACCTTCC	1470
QY	1331	TTCCTTCCCTTAGGGGTGCAGAGGCGCTGTCTCCACAGCTCTGTGTGGGGGTGGAGAGACTCTCT	1390
Db	1471	TTCCTTCCCTTAGGGGTGCAGAGGCGCTGTCTCCACAGCTCTGTGTGGGGGTGGAGAGACTCTCT	1530
QY	1391	GGAGAGAGGTGAGAAGGTGAGAGCTATGCTTTGGGGGGGACAGGCGCAGAAATGAAGTCTGG	1450
Db	1531	GGAGAGAGGTGAGAAGGTGAGAGCTATGCTTTGGGGGGGACAGGCGCAGAAATGAAGTCTGG	1590
QY	1451	GTCAGAGAGCCCAAGCTGGGTCGAGCTGAGCTATGTAAGGCGCTTCAATTGTTGTTCTGAG	1510
Db	1591	GTCAGAGAGCCCAAGCTGGGTCGAGCTGAGCTATGTAAGGCGCTTCAATTGTTGTTCTGAG	1650
QY	1511	ACCCCCACCCCAACGACCAATCCAAATTAAGTGCATCTCCAGCTGAAAAA	1570
Db	1651	ACCCCCACCCCAACGACCAATCCAAATTAAGTGCATCTCCAGCTGAAAAA	1710
QY	1571	AAA 1573	
Db	1711	AAA 1713	
RESULT 6			
LOCUS	AF266756	1693 bp	mRNA linear PRI 01-JUN-2000
DEFINITION	Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.		
VERSION	AF266756		
KEYWORDS	AF266756.1 GI:8133099		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.		
TITLE	1 (bases 1 to 1693)		
	Melendez-A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.		
	Human Sphingosine Kinase, Molecular Cloning, Functional		
	Characterisation and Tissue Distribution		
JOURNAL	Gene (2000) In press		

AUTHORS		2 (bases 1 to 1693)	
TITLE		Melendez,A.J., Carlos-Dias,E., Gostink,M., Allen,J.M. and Takacs,L.	
JOURNAL		Direct Submission Submitted (11-MAY-2000) Department of Molecular and Cellular Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes 94265, France	
FEATURES	Location/Qualifiers		
source	1..1693		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
	/map="17q25.2"		
gene	1..1693		
	/gene="SPHK1"		
	270..1424		
	/gene="SPHK1"		
	/codon_start=1		
	/product="sphingosine kinase"		
	/protein_id="AAF3470.1"		
	/db_xref="GI:8133100"		
	/translation="MDPAGPRGVLPKPCRYVLLINPRGKGLOLFESHVPLAE AELSEFTLMTERNHARELVRSEELGEMDLVYNSGGLHNEVYNGMEPDETAIQ KPLCSLPAGSGNLAALSNHYAGEVQNTDILNCLILCRLLSPNNLLSHTASG LRPLSVLSIAGMFLADVDLHSEKRRRIGEMRFITGFLRLAIRTYRGRLAYLPVGRV GSKTPASPVYVQCGPVDHALVPLEEPPSHMTYVPPDEDFVLVALLSHSGSEFAAP MGRAGVNMFLFVYRAGVSRAMLRLFLFLAEKSGHMYECQYLVYVVAFLRIEKPDSG KGFAVDGELMLVSEAVOQVHPNYFMVYSGCVPEPPSMKDDQMPPEEPL"		
polyA-signal	1676..1681		
	/gene="SPHK1"		
BASE COUNT	279 a 529 c 556 g 329 t		
ORIGIN			
Query Match	97.1%; Score 1527; DB 9; Length 1693;		
Best Local Similarity	99.4%; Pred. No. 1.9e-274;		
Matches 1533; Conservative	0; Mismatches 10; Indels 0; Gaps 0;		
QY	11 CGGCCCTCGCAACGCCCGCCTGGGACAGCACCCGATAGAGAGCTGAAAGGACGAGAGCCGCC	70	
DB	151 CGGCCCTCGCAACGCCCGCCTGGGACAGCACCCGATAGAGAGCTGAAAGGACGAGAGCCGCC	210	
QY	71 ACGGGACAGCGCCCAACAGAGGCGACGGGACCCCTGGACAGGGGAGCGCGGGCGAGGTTA	130	
DB	211 ACGGACAGCGCCCAACAGAGGCGACGGGACCCCTGGACAGGGGAGCGCGGGCGAGGTTA	270	
QY	131 TGGATCCAGCGGGGGGGCCCCCGGGGGCGTGCATCCCGCGGCCCTGCGCGTGCATGCTGTC	190	
DB	271 TGGATCCAGCGGGGGGGCCCCCGGGGGCGTGCATCCCGCGGCCCTGCGCGTGCATGCTGTC	330	
QY	191 TGAACCCCGCGGGGGGACAGGAGGCAAGGCTTTCGAGCTTTCGCGAGTACGTGACGCC	250	
DB	331 TGAACCCCGCGGGGGGACAGGAGGCAAGGCTTTCGAGCTTTCGCGAGTACGTGACGCC	390	
QY	251 TTTTGGCGGAGGCGTGAATTCCTTCACGCGATGCTACAGAGGCGGGGAAACACGCGGC	310	
DB	391 TTTTGGCGGAGGCGTGAATTCCTTCACGCGATGCTACAGAGGCGGGGAAACACGCGGC	450	
QY	311 GGGAGCTGTCGCGTCCGAGGAGAGCTGGGGCCCGCTGGGAGCGTCTGGTGGTCAATGCTGGAG	370	
DB	451 GGGAGCTGTCGCGTCCGAGGAGAGCTGGGGCCCGCTGGGAGCGTCTGGTGGTCAATGCTGGAG	510	
QY	371 ACGGCTGATGACACAGAGTGTGAACGGGCTCATGAGCGGCGCTGACCTGAGAGACCGCA	430	
DB	511 ACGGCTGATGACACAGAGTGTGAACGGGCTCATGAGCGGCGCTGACCTGAGAGACCGCA	570	
QY	431 TCCGAAACCCCTGCTGTAGCCCTCCAGAGGCTCTGGGAAAGCGGTGGACGCTTCCTTGA	490	
DB	571 TCCGAAACCCCTGCTGTAGCCCTCCAGAGGCTCTGGGAAAGCGGTGGACGCTTCCTTGA	630	
QY	491 ACCATTATGCTGCTATGACAGAGTCAACCAATGAGAGCTCCTGACCAACTGACAGCTAT	550	
DB	631 ACCATTATGCTGCTATGACAGAGTCAACCAATGAGAGCTCCTGACCAACTGACAGCTAT	690	
QY	551 TGCTGTGCGCGCGCTGCTGACCCATGAACTGCTGCTGTGACACAGGCTTCGGGGG	610	

D	691	TGCTTGCGCGCGGCTGCTGCTACCCATTAACGTCGTCGTGCAACGCGCTTGCGGC	750
Q	611	TGCGCTTCTCTCTGCTGCTCAGCCTCGCCGTGGGGCTTCAATTGCTGATGTGAGCATAGAGA	670
D	751	TGGCGCTCTTCTCTGTGCTCAGCCTCGCCGTGGGGCTTCAATTGCTGATGTGAGCATAGAGA	810
Q	671	GTCAGAGATATGGCGTCTGGGGAGATACGCTTCACTCTCTGAGGACACTTCTCGGCTGAG	730
D	811	GTCAGAGATATGGCGTCTGGGGAGATATCGGCTTCACTCTGAGGACACTTCTCGGCTGAG	870
Q	731	CACCCCTGCGACCTACGCGGCGGCGACGTGGCTACGTCCTCTGATAGGAACAGTGGTTCCA	790
D	871	CACCCCTGCGACCTACGCGGCGGCGACGTGGCTACGTCCTCTGATAGGAACAGTGGTTCCA	930
Q	791	AGACACATGCCCCCGGCTGTGCTGCACAGAGGCGCGGTATGATGACACCTGTGGCAC	850
D	931	AGACACATGCCCCCGGCTGTGCTGCACAGAGGCGCGGTATGATGACACCTGTGGCAC	990
Q	851	TGAGAGAGCATGCTGCTCTCACTGGACAGTGTGTCCGACAGAGACTTGTGTCTAGTCC	910
D	991	TGAGAGAGCATGCTGCTCTCACTGGACAGTGTGTCCGACAGAGACTTGTGTCTAGTCC	1050
Q	911	TGGCACTGCTGCACTCCGACCTGGCGAGTGAATGTTGCTGCACCCATGGCGCGCTGTG	970
D	1051	TGGCACTGCTGCACTCCGACCTGGCGAGTGAATGTTGCTGCACCCATGGCGCGCTGTG	1110
Q	971	CACCTGGCGCATGCATCTGTCTTACGTCGGGGCGGAGTGTCTGCTGCACCTGCTGTC	1030
D	1111	CACCTGGCGCATGCATCTGTCTTACGTCGGGGCGGAGTGTCTGCTGCACCTGCTGTC	1170
Q	1031	GCCCTTCCCTGGCCATGGAGAGGGGAGGATATGAGATATGATGCCCTACTTGTGAT	1090
D	1171	GCCCTTCCCTGGCCATGGAGAGGGGAGGATATGAGATATGATGCCCTACTTGTGAT	1230
Q	1091	ATCTGCCGTGGTGGCCCTTCCGCTTGAGACCAGACAGATGGGAAGGTGTGTTGACATCG	1150
D	1231	ATCTGCCGTGGTGGCCCTTCCGCTTGAGACCAGACAGATGGGAAGGTGTGTTGACATCG	1290
Q	1151	ATGGGGAATATGATGTTAGCGAGGCGGTCAGGGCGAGGTGTGCAACCCAACTACTTCTGGA	1210
D	1291	ATGGGGAATATGATGTTAGCGAGGCGGTCAGGGCGAGGTGTGCAACCCAACTACTTCTGGA	1350
Q	1211	TGCTCAGCGTGTGCTGAGACCCCGCCGAGCTGGAAGCCCGACGAGATGCGACGCGCAG	1270
D	1351	TGCTCAGCGTGTGCTGAGACCCCGCCGAGCTGGAAGCCCGACGAGATGCGACGCGCAG	1410
Q	1271	AAGAGCCCTTATGACCCCTGGGGCGCGCTGTGCCCTGATCTTACTTGTGAGGACCCCTTC	1330
D	1411	AAGAGCCCTTATGACCCCTGGGGCGCGCTGTGCCCTGATCTTACTTGTGAGGACCCCTTC	1470
Q	1331	TCTCTCCCTAGGGCTGGAAGGGCTGTGTCCACAGTCCTGTGGGGGTGGAGAGACCTCT	1390
D	1471	TCTCTCCCTAGGGCTGGAAGGGCTGTGTCCACAGTCCTGTGGGGGTGGAGAGACCTCT	1530
Q	1391	GGAAGAGGTGAGAGGTGAGAGGCTATGTTTGGGGGACAGAGCCAGAAATGAAGTCTCG	1450
D	1531	GGAAGAGGTGAGAGGTGAGAGGCTATGTTTGGGGGACAGAGCCAGAAATGAAGTCTCG	1590
Q	1451	GTCAGAGGCCAGCTGGCTGGGCGCCAGCTGCCATGTGAAGGCTTCTAGTTTGTCTGAC	1510
D	1591	GTCAGAGGCCAGCTGGCTGGGCGCCAGCTGCCATGTGAAGGCTTCTAGTTTGTCTGAC	1650
Q	1511	ACCCCGACCCCGACACCAATTCAAATTAAGTAGACTTCCCA	1553
D	1651	ACCCCGACCCCGACACCAATTCAAATTAAGTAGACTTCCCA	1693

RESULT 7			
AX224379			
LOCUS	AX224379	1600 bp	DNA
DEFINITION	Sequence		1 linear
ACCESSION	AX224379	1 from Patent WO0160990.	PAT 10-SEP-2001

VERSION	AX224379.1	GI:15554031
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
REFERENCE	1 (bases 1 to 1600)	
AUTHORS	Rastelli, L.	
TITLE	Novel sphingosine kinases	
JOURNAL	Patent: WO 0160990-A.1. 23-Aug-2001; Curagen Corporation (US); GENENTECH, INC. (US)	
FEATURES	Location/Qualifiers	
Source	1..1600	
BASE COUNT	265 a 492 c 331 g 311 t	1 others
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"	
Query Match	96.1%; Score 1511.6; DB 6; Length 1600;	
Best Local Similarity	98.7%; P: 0; Mismatches 19; Indels 1; Gaps 1;	
Matches 1534: Conservative		
QY	1 GCCCCACAGCCGGCCCTGCACACCCCGCCTGGGAGCAGCACCAGATTAAGAGACTGAAGGCAGG 60	
Db	46 GCCCCACAGCCGGCCCTGCACACCCCGCCTGGGAGCAGCACCAGATTAAGAGACTGAAGGCAGG 105	
QY	61 AGCCGCGCCACAGGGGAGCG-CCGCCAGGCCCAAGGGAGCCCTCTGGCAGCGAGCCGCG 119	
Db	106 AGCCGCGCCACAGGGGAGCGCCGCCCAAGGCCCAAGGGAGCCCTCTGGCAGCGAGCCGCG 165	
QY	120 GCGTCGAGGTTATGATCCAGCGCGCGCCCGCGGGGGTGTCTCCCGCGCCCTGCGCGCT 179	
Db	166 GGTGAGGTTATGATCCAGCGGGGGGCGCCCGGGGGGTGTCTCCCGCGCCCTGCGCGCT 225	
QY	180 GCGTGGGTGTGTAACCCGCGCGCGGCAAGGGGAAAGGGCTTCAGACTCTTCGCGAGTCA 239	
Db	226 GCTGGTGTGTGTAACCCGCGCGCGGCAAGGGGAAAGGGCTTCAGACTCTTCGCGAGTCA 285	
QY	240 GGTGAGCGCCCTTTGGCTGAGGCTGAATCGCTTCAGGCTGATCTCAGTGAAGCGCG 299	
Db	286 GGTGAGCGCCCTTTGGCTGAGGCTGAATCGCTTCAGGCTGATCTCAGTGAAGCGCG 345	
QY	300 GAACACGCGCGGGAGCTGCTGCGCTCGAGAGAGCTTGGGCGCTGGCAGCTCTGTGTGT 359	
Db	346 GAACACGCGCGGGAGCTGCTGCGCTCGAGAGAGCTTGGGCGCTGGGAGCTCTGTGTGT 405	
QY	360 CATGCTGTGAGACGGGCTGATGACAGAGTGATGAACGGGCTCATATGAGAGCGACTG 419	
Db	406 CATGCTGTGAGACGGGCTGATGACAGAGTGATGAACGGGCTCATATGAGAGCGACTG 465	
QY	420 GGAGAGCGCCATCCAGAGCCCGCTGTGTAGAGCTCCCAAGCAGGCTCTGGCAGCGCGTGG 479	
Db	466 GGAGAGCGCCATCCAGAGCCCGCTGTGTAGAGCTCCCAAGCAGGCTCTGGCAGCGCGTGG 525	
QY	480 AGCTTCCTTGAAACCATATGCTGGCTATGAGACAGCTACCAATGAAGAAGCTCTGACCAA 539	
Db	526 AGCTTCCTTGAAACCATATGCTGGCTATGAGACAGCTACCAATGAAGAAGCTCTGACCAA 585	
QY	540 CTGACAGGCTATGCTGTGGCGCGGCGGTGCTGTCAGACCCCAATCAACCTTCAGTCTGACAC 599	
Db	586 CTGACAGGCTATGCTGTGGCGCGGCGGTGCTGTCAGACCCCAATCAACCTTCAGTCTGACAC 645	
QY	600 GGGCTTCGGGGGTGGCCCT 659	
Db	646 GGGCTTCGGGGGTGGCCCT 705	
QY	660 GGAACCTAAGAGTGAAGATATCGCGCTCTGGGGAGATGCGCTTCACTCTGGGACCTT 719	
Db	706 GGAACCTAAGAGTGAATAGATGCGGCTCTGGGGAGATGCGCTTCACTCTGGGACCTT 765	
QY	720 CCGTGCCTGTGGCAACCTCTGGCAACCTACCGCGGCGAGCTGGCTTCACTCTCCGTAGGAAG 779	
Db	766 CCGTGCCTGTGGCAACCTCTGGCAACCTACCGCGGCGAGCTGGCTTCACTCTCCGTAGGAAG 825	

```

QY 780 AGTGGTTCAGAGACACCTGCTCCCGGTTGTGTCCAGAGAGGCGCGATGATGACA 839
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 826 AGTGGTTCAGAGACACCTGCTCCCGGTTGTGTCCAGAGAGGCGCGATGATGACA 885
QY 840 CCTGTGTGACATGAGAGAGGACCAATGCTCTGATGACAGTGTGCGGAGAGACTT 899
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 886 CCGGTGTGACATGAGAGAGGACCAATGCTCTGATGACAGTGTGCGGAGAGACTT 945
QY 900 TGTGTGTGACATGAGAGAGGACCAATGCTCTGATGACAGTGTGCGGAGAGACTT 959
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 946 TGTGTGTGACATGAGAGAGGACCAATGCTCTGATGACAGTGTGCGGAGAGACTT 1005
QY 960 GGGCGGCTGTGAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1019
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1006 GGGCGGCTGTGAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1065
QY 1020 CATGTGTGTGAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1079
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1066 CATGTGTGTGAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1125
QY 1080 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1126 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1185
QY 1140 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1186 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1245
QY 1200 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1246 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1305
QY 1260 GCGACCGGCAAGAGAGGCTTATGATGATGATGATGATGATGATGATGATGATG 1319
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1306 GCGACCGGCAAGAGAGGCTTATGATGATGATGATGATGATGATGATGATGATG 1365
QY 1320 AGGACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1379
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1366 AGGACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1425
QY 1380 GAGACTCTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1439
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1426 GAGACTCTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1485
QY 1440 TGAAGTCTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1499
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1486 TGAAGTCTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1545
QY 1500 TTTGTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1553
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1546 TTTGTCTGAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1599

RESULT 8
LOCUS BC009419 1869 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:15040 IMAGE:3506924, mRNA, complete cds.
ACCESSION BC009419
VERSION BC009419.1 GI:14495624
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1869)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

```

```

COMMENT
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fiedel, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Jilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stolt,
Michael Thorne, Miranada Teal, Natalja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Zaira.

FEATURES
Source
Location/Qualifiers
1..1869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15040 IMAGE:3506924"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_1db="NIH_MGC_14"
/1db_host="DH10B-R"
/note="vector: pOTB7"
381..1577
/codon_start=1
/product="unknown (protein for MGC:15040)"
/protein_1d="AAH09419.1"
/db_xref="GI:14495625"
/translation="MDPVGCGRGLFGFVSAGPRGVLPDPCRYLVLNPRGKGA
LQLFPSHVPPLAEISFTLMEERHAEIVSEELGEMDALVYVSGGLHNEV
NLGMRPMEPAIOGLPISLPAAGSNALAEINHYAGEVQYNTNDLLNCLLLRL
LSPMLNLSLHTASGLRLFSVSLAMGPIADVDLSEKRRIRGEBRFTLCTLRALALR
TYRGLATLPGVRGSKTPASPVVYVQDQPVANHLVPLEPVSHTVVPDDEFLVLA
LHLSHLSSEMPAAMPGRCAAGVMPHLEVRABVSRLMLRLFLAMEKGRHMEVECPYL
YVPVVAERLEPRKDGKGYFAYVDGELMVSFAVGQVHPNFYVMSGVEPSPMKPQOMP
PPEPLI"

BASE COUNT 322 a 570 c 623 g 354 t
ORIGIN
Query Match 95.8%; Score 1506.8; DB 9; Length 1869;
Best Local Similarity 97.2%; Pred. No. 1.1e-270;
Matches 1571; Conservative 0; Mismatches 2; Indels 43; Gaps 2;
QY 1 GCGCCACAGCCGCGCTGAGAGCGCCGCTGGGAGACACCGATAGAGACTGAGGAGG 60
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 251 GCGCCACAGCCGCGCTGAGAGCGCCGCTGGGAGACACCGATAGAGACTGAGGAGG 310
QY 61 AGCCGCGCGCAGGAGAGG-CCCCACAGCGCAGAGGACCCCTGACAGCGGAGCGGG 119
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 311 AGCCGCGCGCAGGAGAGGCGCCGCCACAGCGCAGGAGACCCCTGACAGCGGAGCGGG 370
QY 120 GGTGAGGTTTATGATC-----C 137
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 371 GGTGAGGTTTATGATGATCAGTGTGCGGAGCTGCGCTCTTTGGTTTGTTC 430
QY 138 AGCGGGGCGCGCGCGGCGTGTCTCCGCGGCGCTGCGCTGTGTGTGTGACACC 197
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 431 AGCGGGGCGCGCGCGGCGTGTCTCCGCGGCGCTGCGCTGTGTGTGTGACACC 490
QY 198 GCGCGCGCAGAGAGGAGGCGCTTGTGAGCTTTCGAGAGTCAGTACAGCCCTTTTGGC 257
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 491 GCGCGGCGCAGAGAGGCGCTTGTGAGCTTTCGAGAGTCAGTACAGCCCTTTTGGC 550
QY 258 TGAGGCTGAATCTCCTTACAGCTGATGCTGAGCGGCGGAGACACAGCGGCGGAGCT 317
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```


Db	551	TTGAGCGTGA	AAATCTCCTCTTACAGCTGATGCTACATGAGACCGGAAACACGCGGGAGCCT	610
Oy	318	GGTGGCGTGG	AGAGAGCTGGGGCCGCTGGAGCGCTGTGGTGGATATGTCTGSAGACGGGCT	377
Db	611	GGTGGCGTGG	AGAGAGCTGGGGCCGCTGGAGCGCTGTGGTGGATATGTCTGSAGACGGGCT	670
Oy	378	GATGACAGAG	GGTGGTGAACGGGGCTCATGGACGGGCTGATCTGGAGACCGCCCTATCCAGA	437
Db	671	GATGACAGAG	GGTGGTGAACGGGGCTCATGGAGCGGCTGATCTGGAGACCGCCCTATCCAGA	730
Oy	438	GGCCCGTGT	ACCCCTCCAGAGGCTTGGCAACGGCTGGGACCTTCTTGAACCATTA	497
Db	731	GGCCCGTGT	ACCCCTCCAGAGGCTTGGCAACGGGCTGGGAGGCTTCTTGAACCATTA	790
Oy	498	TGCTGGCTAT	GACAGGTCAACCAATGAGACCTCTGACCAACTGCACGCTATTGTCTGTG	557
Db	791	TGCTGGCTAT	GACAGGTCAACCAATGAGACCTCTGACCAACGCTTCTGGGCTGTGTG	850
Oy	558	CGCGCGGCTG	CTGTACCCACTGAACCTGCTGTCTGTGCACAGGCTTCTGGGCTGTGGCT	617
Db	851	CGCGCGGCTG	CTGTACCCACTGAACCTGCTGTCTGTGCACAGGCTTCTGGGCTGTGGCT	910
Oy	618	CTTCTGTGTG	CTCAAGCTGGAGCTGGAGCTTCAATGTGATGTGACCTAAGAAGTGAAGA	677
Db	911	CTTCTGTGTG	CTCAAGCTGGAGCTGGAGCTTCAATGTGATGTGACCTAAGAAGTGAAGA	970
Oy	678	GATTCGGGCT	CTGGGGAGATGGCGCTTCACTGTGGGACACTTTCGCGTGTGGACGCTT	737
Db	971	GATTCGGGCT	CTGGGGAGATGGCGCTTCACTGTGGGACACTTTCGCGTGTGGACGCTT	1030
Oy	738	GGCCACACTA	CGCCGGCGCACTGGGCTACGCTCCGTGAGGAAGAGTGGGTTTCCAAAGAC	797
Db	1031	GGCCACACTA	CGCCGGCGCACTGGGCTACGCTCCGTGAGGAAGAGTGGGTTTCCAAAGAC	1090
Oy	798	TGGCTCCCCG	TTGTGTGTCACAGAGGGCCGGTAAATGCACACCTTGTGTGCATGTGAGA	857
Db	1091	TGGCTCCCCG	TTGTGTGTCACAGAGGGCCGGTAAATGCACACCTTGTGTGCATGTGAGA	1150
Oy	858	GCGAGTGGCC	CTCTCACTGAGAGTGGTCCGACAGAGACTTGTGTGTACTCTGGGACT	917
Db	1151	GCGAGTGGCC	CTCTCACTGAGAGTGGTCCGACAGAGACTTGTGTGTACTCTGGGACT	1210
Oy	918	GCTGTCACTG	GCACCTGGGCAATGATGTTTGTGTGACCCATGGGCGCTGTGACGTGG	977
Db	1211	GCTGTCACTG	GCACCTGGGCAATGATGTTTGTGTGACCCATGGGCGCTGTGACGTGG	1270
Oy	978	CGTCATGCAT	CTGTCTTACTACGTGGGCGGAGAGTGTGTGTGCATGTGTGCGCCCTTT	1037
Db	1271	CGTCATGCAT	CTGTCTTACTACGTGGGCGGAGAGTGTGTGTGCATGTGTGCGCCCTTT	1330
Oy	1038	CGTGGGCATG	AGAGAGGCAAGGATATGAGATATAAAGCCCTACTTGTATATGTGC	1097
Db	1331	CGTGGGCATG	AGAGAGGCAAGGATATGAGATATAAAGCCCTACTTGTATATGTGC	1390
Oy	1098	CGTGGTCCG	CTCTCCGCTTGGAGCCCAAGGATGGGAAAGTGTGTTTCACTGATGGGA	1157
Db	1391	CGTGGTCCG	CTCTCCGCTTGGAGCCCAAGGATGGGAAAGTGTGTTTCACTGATGGGA	1450
Oy	1158	ATTGATGTTT	ACGAGCGCGTGAAGGCGGACAGGTGCACCCCAATCACTTGTGATGGTGTG	1217
Db	1451	ATTGATGTTT	ACGAGCGCGTGAAGGCGGCGGAGGCGCCAGGTGCACCCCAATCACTTGTGATGGTGTG	1510
Oy	1218	CGGTTTCGTA	GAGCCCGCCGCAAGTGAAGGCCAGAGATGTGCACCGGCAGAAAGAGC	1277
Db	1511	CGGTTTCGTA	GAGCCCGCCGCCCGCCGCAAGTGAAGGCCAGAGATGTGCACCGGCAGAAAGAGC	1570
Oy	1278	CTTATGACCC	CTGGGCGGCTGTGCTTAAAGTGTCTACTTGCAGAGACCTTCTCTTTC	1337
Db	1571	CTTATGACCC	CTGGGCGGCTGTGCTTAAAGTGTCTACTTGCAGAGACCTTCTCTTTC	1630
Oy	1338	CTAAGGCTTG	CAGAGGCTGTGCACAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAG	1397
Db	1631	CTAAGGCTTG	CAGAGGCTGTGTCTCCACAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAG	1690

[illegible]

Db 1145 GCGGCGGCGGCAAGGCGAAGGCGCTTGCAGCTCTCCGAGTCACTGCAAGCCCTTTTGG 1204
Qy 257 CTGAGGCTGAATTCCTTTCACGCTGATGCTCAGTGAAGGCGGCAACCAACGCGGAGAC 316
Db 1205 CTGAGGCTGAATTCCTTTCACGCTGATGCTCAGTGAAGGCGGCAACCAACGCGGAGAC 1264
Qy 317 TGGTCCGCTCGGAGAGCTGCGGCGCTGGGACGCTCTGGTGGTCACTGCTGAGAGAGGCG 376
Db 1265 TGGTCCGCTCGGAGAGCTGCGGCGCTGGGAGCGCTCTGGTGGTCACTGCTGAGAGAGGCG 1324
Qy 377 TGATGAGAGAGTGGTGAACCGGCTCATGAGAGCGGCTCAGTGGGAGAGCGGCTTCACA 436
Db 1325 TGATGAGAGAGTGGTGAACCGGCTCATGAGAGCGGCTCAGTGGGAGAGCGGCTTCACA 1384
Qy 437 AGCCCTGTGTAGCTCCAGAGAGCTGGAACGCGCTGACAGCTTCTTGAACCAAT 496
Db 1385 AGCCCTGTGTAGCTCCAGAGAGCTGGAACGCGCTGACAGCTTCTTGAACCAAT 1444
Qy 497 ATGCTGGCTATGAGAGAGTGAACCAATGAAGACCTCTGACCACTGACAGCTATTGCTGT 556
Db 1445 ATGCTGGCTATGAGAGAGTGAACCAATGAAGACCTCTGACCACTGACAGCTATTGCTGT 1504
Qy 557 GCGGCGGCTGCTGTCACCCATGAACCTGCTGTCTGACACGCGCTTGGGCGCTGCGC 616
Db 1505 GCGGCGGCTGCTGTCACCCATGAACCTGCTGTCTGACACGCGCTTGGGCGCTGCGC 1564
Qy 617 TCTTCTGTGTGCTAGACCTGCGGCTGAGGCTTCACTGATGAGAGAGTGAAGAGTGA 676
Db 1565 TCTTCTGTGTGCTAGACCTGCGGCTGAGGCTTCACTGATGAGAGAGTGAAGAGTGA 1624
Qy 677 AGTATGCGGCTGCGGAGAGTGAACCTGCTTCACTGAGGACCTTTCGCTGTCGACGCC 736
Db 1625 AGTATGCGGCTGCGGAGAGTGAACCTGCTTCACTGAGGACCTTTCGCTGTCGACGCC 1684
Qy 737 TGGGACACCTACCGGCGGCGGCTGCTGCTTCTGAGAGAGTGGTTCACAGACAC 796
Db 1685 TGGGACACCTACCGGCGGCGGCTGCTGCTTCTGAGAGAGTGGTTCACAGACAC 1744
Qy 797 CGGCTCCCGCTTGTGCTGACAGAGGCGGCTGATGATGACACCTTGTGCCACTGGAAG 856
Db 1745 CGGCTCCCGCTTGTGCTGACAGAGGCGGCTGATGATGACACCTTGTGCCACTGGAAG 1804
Qy 857 AGCCAGTGGCTCTCAGTGAAGTGTGCGGAGAGGAGCTTGTGCTAGTCTGAGAC 916
Db 1805 AGCCAGTGGCTCTCAGTGAAGTGTGCGGAGAGGAGCTTGTGCTAGTCTGAGAC 1864
Qy 917 TGGTGCACCTGCGAGCTGGGAGATGATGCTGCTGACACCATGGCGCTGTGACGCTG 976
Db 1865 TGGTGCACCTGCGAGCTGGGAGATGATGCTGCTGACACCATGGCGCTGTGACGCTG 1924
Qy 977 GCGTATGATGCTGTCTCAGTGGGCGGAGTGTCTGCTGACATGCTGCGGCTCT 1036
Db 1925 GCGTATGATGCTGTCTCAGTGGGCGGAGTGTCTGCTGACATGCTGCGGCTCT 1984
Qy 1037 TCCCTGGCCATGAGAGAGGCGAGCATATGAGATGAGCCCTACTTGTATATATGCTG 1096
Db 1985 TCCCTGGCCATGAGAGAGGCGAGCATATGAGATGAGCCCTACTTGTATATGCTG 2044
Qy 1097 CCGTGGTGGCTTCCGCTTGGAGCCAGAGATGGAAGAGTGTGCTGAGTGGAGTGGG 1156
Db 2045 CCGTGGTGGCTTCCGCTTGGAGCCAGAGATGGAAGAGTGTGCTGAGTGGAGTGGG 2104
Qy 1157 AATTGATGTTGAGAGGCGGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1216
Db 2105 AATTGATGTTGAGAGGCGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2164
Qy 1217 GCGGTGGCTGAGAGGCGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1276
Db 2165 GCGGTGGCTGAGAGGCGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2224
Qy 1277 CCTATGAGACCTTGGGCGGCTGCTGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTG 1336
Db 2225 CCTATGAGACCTTGGGCGGCTGCTGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTG 2284

Qy 1337 CCTAGGGCTGACAGGCGCTTCCACAGCTCTGTTGGGGTGAAGAGACTTCTTGGAGAA 1396
Db 2285 CCTAGGGCTGACAGGCGCTTCCACAGCTCTGTTGGGGTGAAGAGACTTCTTGGAGAA 2344
Qy 1397 GGGTGAAGGTGAGAGGCTTGGGCGGAGAGGCGGAGAGTGAAGTCTGGTGAAG 1456
Db 2345 GGGTGAAGGTGAGAGGCTTGGGCGGAGAGGCGGAGAGTGAAGTCTGGTGAAG 2404
Qy 1457 AGCCGAGTGGCTGGGCGGAGAGTCTATGAGGCTTCTATTTGTTGAGACCCC 1516
Db 2405 AGCCGAGTGGCTGGGCGGAGAGTCTATGAGGCTTCTATTTGTTGAGACCCC 2464
Qy 1517 ACCCGAGAGCAAAATCAAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1554
Db 2465 ACCCGAGAGCAAAATCAAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2502

RESULT 10
HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
LOCUS HSA245504
DEFINITION Homo sapiens mRNA for sphingosine kinase (gene SPK).
ACCESSION AJ245504
VERSION AJ245504.1 GI:8017375
KEYWORDS sphingosine kinase; SPK gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Van Veldhoven, P. P. and Gijssels, S.
JOURNAL Unpublished
TITLES 2 (bases 1 to 1428)
AUTHORS Van Veldhoven, P. P.
DIRECT SUBMISSION
SUBMITTED (16-AUG-1999) Van Veldhoven P. P., Molecular Cell Biology,
katholieke Universiteit Leuven, K.U.Leuven, Campus Gasthuisberg,
Afd. Farmakologie, Herestraat, B-3000 Leuven, BELGIUM
REMARK Revised by author (17-MAY-2000)
FEATURES
location/Qualifiers
source 1..1428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24"
1..1146
/gene="SPK"
1..1146
/gene="SPK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="CAB92131.1"
/db_xref="GI:8017375"
/translation="MDPAGPCVARLRLVLLNPRGKAKALQFRSHVOPILAEAI
SPTLUTERIRNIRARELSEELGRMDALVYMGDDGIMHEVNLGRPMETAIKPL
CSLPAGSGKPLAASLNHYAGVEOVNEDILTCMTLLICRLISPNLLISLHVASGLRL
PSVLSLAKETINDVLESERKRLDEMRITLCTFLRLALRLRYRGLATLPLGRGSK
TPASPVVQOQVDHLVPLEPVPSSHVVVDEDEVLVLLHSLGSEMPAEPGR
CAAGVHLFVYRAGVSRAMLRLFLAMEGRIMEXECPYLIVYPVGAFLPEKDKGV
FAVDELWSEVVOQVHPNYPFMVSGCEVPPSPKPOQMPPEEPL"
BASE COUNT 247 a 422 c 462 g 297 t
ORIGIN

Query Match 88.1%; Score 1385.6; DB 9; Length 1428;
Best Local Similarity 99.7%; Pred. No. 4.1e-248;
Matches 1388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 175 GCGGTGGCTGCTGAGAACCGGCGGAGGAGGCAAGGCGGCTTGAAGTCTCCGCG 234
Db 37 GCGGTGGCTGCTGAGAACCGGCGGAGGAGGCAAGGCGGCTTGAAGTCTCCGCG 96
Qy 235 ACTGACGTGAGCCCTTTGGCTGAGAGCTGAAGTCTCTTACGCTGATGCTGAGTGA 294

[illegible]

```

/db_xref="GI:9280106"
/translation="MDPAGPBGVLPSCRCHVLLINRGGKGLQLFERSHVDPLAE
AEISFTLMEERHRAELVRSBELGMDLIVYSGGMLHVEVNGMERDDETATQ
KPLCSLPAGSGNLAASLAINVAGYEYDNDILJNCRILCRLLISPMNLSTFASG
LRLEPVSILAMGFLADYDLSEKBRIGENRFTLGTFLRLAALTLLKRIALFVGR
GSKTPVSVVYVGGVPDAHVPLEEPVPSMTVTVVDTFVLVLLSHLSEFPAF
MGRRAAGVMHLFVIRACVSRAMLRLFLAMEKGRHMEYECPELVYISGWDIG"
BASE COUNT      355 a      639 c      650 g      371 t
ORIGIN

Query Match      83.1%; Score 1306.6; DB 9; Length 2015;
Best Local Similarity 92.3%; Pred. No. 2e-233;
Matches 1452; Conservative 0; Mismatches 64; Indels 57; Gaps 5;

Oy 1 GCGCCACAGCGCGCGCTCGAGACCGCGCTGGGCGAGCACCGCTAAGAGAGCTGAAGCAGG 60
Db 498 GCGCCACAGCGCGCGCTCGAGACCGCGCTGGGCGAGCACCGCTAAGAGAGCTGAAGCAGG 557
Oy 61 AGCGCGCCGACGAGCGGCGAGCGCCCG-ACAGCGCCAGAGGACCGCCCTGGCAGCGGAGCGCG 119
Db 558 AGCGCGCCGACGAGCGGCGAGCGCCCGAGAGCGGAGGAGCGCCCGCGAGAGGAGCGAGC 617
Oy 120 GGTGAGGTATGATCCAGCGGCGGCGCGCGCGCGCTGCTCCCGCGCGCTGCCGCGT 179
Db 618 GGTGAGGTATGATCCAGCGGCGGCGCGCGCGCGCTGCTCCCGCGCGCTGCCGCGT 677
Oy 180 GCTGGTCTGTGTAACCGCGCGCGCGCGAGAGGCGCTTGCAGCTCTCCGAGTCA 239
Db 678 GCTGGTCTGTGTAACCGCGCGCGCGCGAGAGGCGCTTGCAGCTCTCCGAGTCA 737
Oy 240 CGTGAGCGCGCTTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTACTAGACGCGC 299
Db 738 CGTGAGCGCGCTTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTACTAGACGCGC 797
Oy 300 GAGCAGCGCGGAGAGCTGTGCGCTCGAGAGAGCTGGGCGCGCTGGAGCGCTCTGTGCT 359
Db 798 GAGCAGCGCGGAGAGCTGTGCGCTCGAGAGAGCTGGGCGCGCTGGAGCGCTCTGTGCT 857
Oy 360 CATGCTGTGAGAGCGGCTGATGCAAGAGTGGTGAAGGCGCTCATGAGCGCGCTGACGTG 419
Db 858 CATGCTGTGAGAGCGGCTGATGCAAGAGTGGTGAAGGCGCTCATGAGCGCGCTGACGTG 917
Oy 420 GGAGAGCGCCATCCAGAAAGCCCTGTGTAGCTTCCAGAGGCTCTGGCAACGCGCTGAC 479
Db 918 GGAGAGCGCCATCCAGAAAGCCCTGTGTAGCTTCCAGAGGCTCTGGCAACGCGCTGAC 977
Oy 480 AGCTTCTTGAACCACTTATGCTGCTATGAGAGAGTCAACCAATGAAGACTCTCTGACCA 539
Db 978 AGCTTCTTGAACCACTTATGCTGCTATGAGAGAGTCAACCAATGAAGACTCTCTGACCA 1037
Oy 540 CTGACGCTATTGCTGTGCGCGCGCGCTGCTGACCCATGAACCGTGTCTGTGACAC 599
Db 1038 CTGACGCGGCTGCTGTGCGCGCGCGCTGCTGACCCATGAACCGTGTCTGTGACAC 1097
Oy 600 GCGCTTGGGCGGCTGCGCTCTTCTGTGCTGAGCTTGGGCTTGGGCTTCAATGCTGATGT 659
Db 1098 GCGCTTGGGCGGCTGCGCTCTTCTGTGCTGAGCTTGGGCTTGGGCTTCAATGCTGATGT 1157
Oy 660 GAGCTAAGAGAGTGAAGTATCGGCGCTGTGGGAGAGATGCGCTTCACTGTGGGACCTT 719
Db 1158 GAGCTAAGAGAGTGAAGTATCGGCGCTGTGGGAGAGATGCGCTTCACTGTGGGACCTT 1217
Oy 720 CCGTGGTGTGGAGCGCGGCGAGCTACCGCGGCGGCGAGCTGCGCTACCTCCCTGTAGGAG 779
Db 1218 CCGTGGTGTGGAGCGCGGCGAGCTACCGCGGCGGCGAGCTGCGCTACCTCCCTGTAGGAG 1277
Oy 780 AGTGAGTTCACAGACACCTGCTCCCGCTGTGTGTCTCCAGAGCGGCGCGGTATGATGACA 839
Db 1278 AGCGCGGCTCTAAGACACCTGTTTCCCGCTGTGTGTCTCCAGAGCGGCGCGGTATGATGACA 1337
Oy 840 CCTTGTGCGACATGAGAGGAGGAGGAGTCCCTCTCAGTGAAGAGTGTGCGCGGAGAGACTT 899
Db 1338 CCTTGTGCGACATGAGAGGAGGAGGAGTCCCTCTCAGTGAAGAGTGTGCGCGGAGAGACTT 1397

```

```

Oy 900 TGTGCTAGTCTGTGACCTGCTGACCTGCGACCTGGGCGAGTGAATGTTTGTGACCCAT 959
Db 1398 CTTGCTGTGTGCTAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
Oy 960 GGGCGCGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1019
Db 1458 GGGCGCGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1517
Oy 1020 CATGCTGTGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 1518 CATGCTGTGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1577
Oy 1080 CTACTTGTATATGTGTGCGCGTGTGCGCTTCCGCTTGGAGCCCAAGATGGAAGAGTGT 1139
Db 1578 TTACTTGTATATCAT----- 1592
Oy 1140 GTTTCAGTGTGAGTGGGGAATGATGTTAGCGAGGCGCTGCAAGGCGCAGGTGACCCAAA 1199
Db 1593 -----CAGTGTGAGTGGGGAATGATGTTAGCGAGGCGCTGCAAGGCGCAGGTGACCCAAA 1647
Oy 1200 CTACTTGTGATGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
Db 1648 CTACTTGTGATGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1707
Oy 1260 GCCAGCGCGAGAGAGCCCTTATGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1319
Db 1708 GCCAGCTCCAGAGAGAGCCCTTATGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1767
Oy 1320 AGAGCCCTTCTCTCTTCCCTAGGCGTGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
Db 1768 AGAGCCCTTCTCTCTTCCCTAGGCGTGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826
Oy 1380 GAGACTCTCTGGAAGAAGGTTGAGAGGTTGAGAGCTATGCTTTGGGGGAGAGCGCCAGAA 1439
Db 1827 GAGACTCTCTGGAAGAAGGTTGAGAGGTTGAGAGCTATGCTTTGGGGGAGAGCGCCAGAA 1886
Oy 1440 TGAAGTCTT-GGGTCAAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1498
Db 1887 TGAAGTCTT-GGGTCAAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942
Oy 1498 GTTGTCTGAGAGCGCCCGCCAGCCAGAGCAAAATCAAAATGAAGTGTGCTGCTGCTGCTG 1558
Db 1943 GTTGTCTGAGAGCGCCCGCCAGCCAGAGCAAAATCAAAATGAAGTGTGCTGCTGCTGCTG 2002
Oy 1559 AAAAAAAAAAAAAA 1571
Db 2003 AAAAAAAAAAAAAA 2015

RESULT 12
AF200328 1173 bp mRNA linear PRI 25-AUG-2000
LOCUS AF200328
DEFINITION Homo sapiens sphingosine kinase (SPK) mRNA, complete cds.
ACCESSION AF200328
VERSION AF200328.1 GI:9909360
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1173)
AUTHORS Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A., Xia,P.,
Gamble,J.R., Vadas,W.A. and Wattenberg,B.W.
TITLE Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
BIOCHEM. J. 350 Pt 2, 429-441 (2000)
JOURNAL MEDLINE 20407120
PUBMED 10947957
REFERENCE 2 (bases 1 to 1173)
AUTHORS Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A.B., Xia,P.,
Gamble,J.R., Vadas,W.A. and Wattenberg,B.W.
TITLE Direct Submission

```

JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES Location/Qualifiers
source 1.1173
/organism="Homo sapiens"
/db.xref="taxon:9606"
/cell.type="vascular endothelial"
1.1173
/gene="SPHK"
19.1173
/gene="SPHK"
/codon.start=-1
/product="sphingosine kinase"
/protein_id="AA01980.1"
/db.xref="GI:9909361"
/translation="MDPAGPRGVLPRPCRVLLINPRGKALQLFPSHVOPILLAE
AIEFTLMLTERNNHARELVSEBELRDALVNSGGGLHVEVNLMLERPMETAIQ
KPLCSLPAGSGLALASLNHAYEQVNDLNLCTLLRLSLPMULSLHTASG
LRLFSVLSLWGTADVDLESEKRYRLGEMRFTLGTFLRLAALRTYRGLATLPVGRY
GSKTPASPVVQGGPVDAHLVPLEBPSVSHMTVPDEDFVLVLLVLAH/GSEMPAP
MGRCAAGMHLFVRAGVSRAMLRLFLAMEKGRMEYECPLYLVVPSHAFLEPKDG
KMFVADGELMVESEAVQGVHPNYFMWVSGCEPSPMKDQMPPEEPL"

BASE COUNT 188 a 349 c 390 g 246 t
ORIGIN

Query Match 74.1% Score 1165; DB 9; Length 1173;
Best Local Similarity 99.6% Pred. No. 5e-207;
Matches 1168; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

112 GAGCCGGGGTGGAGTTATGATTCACAGCGGGGCGCCCGGGGGGCTCTCCGGGGCC 171
1 GAGCCGGGGTGGAGTTATGATTCACAGCGGGGCGCCCGGGGGCTCTCCGGGGCC 60
172 TGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 231
61 TGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120
232 CGGAGTCAGCTGACGCCCCCTTTGGCTGAGCTGAATCTCTTACGCTGATGCTACT 291
121 CGGAGTCAGCTGACGCCCCCTTTGGCTGAGCTGAATCTCTTACGCTGATGCTACT 180
292 GAGCGCGGAGACCGACCGCGGAGCTGTCGGTGGAGAGCTGGGCGCTGGAGAGCT 351
181 GAGCGCGGAGACCGACCGCGGAGCTGTCGGTGGAGAGCTGGGCGCTGGAGAGCT 240
352 CTGCTGATGATGCTGAGAGCGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 411
241 CTGCTGATGATGCTGAGAGCGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 300
412 CTGCTGATGATGCTGAGAGCGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 471
301 CTGCTGATGATGCTGAGAGCGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 360
472 GCGCTGGAGCTTCTTGAACCATTAATGCTGATGAGAGTGGTGAACCATTAATGAGAGCTG 531
361 GCGCTGGAGCTTCTTGAACCATTAATGCTGATGAGAGTGGTGAACCATTAATGAGAGCTG 420
532 CTGACCAACTGACAGCTTATGCTGCGCGGCTGCTGACCAACTGACAGCTGCTGCT 591
421 CTGACCAACTGACAGCTTATGCTGCGCGGCTGCTGACCAACTGACAGCTGCTGCT 480
592 CTGACCAACTGACAGCTTATGCTGCGCGGCTGCTGACCAACTGACAGCTGCTGCT 651
481 CTGACCAACTGACAGCTTATGCTGCGCGGCTGCTGACCAACTGACAGCTGCTGCT 540
652 GCGATGTTGAGAGTGAAGTGAAGTATGAGCTGCTGGGGAGAGTCCCTTACTCTG 711
541 GCGATGTTGAGAGTGAAGTGAAGTATGAGCTGCTGGGGAGAGTCCCTTACTCTG 600
712 GCGACCTTCTGCTGGAGAGCTTACAGAGCTTACAGAGCTTACAGAGCTTACAGAGCT 771
601 GCGACCTTCTGCTGGAGAGCTTACAGAGCTTACAGAGCTTACAGAGCTTACAGAGCT 660

772 GTAGAGAGTGGGTTCCAGAGACACCTGCTCCCGGTTGTGTCAGAGAGGCGCCGTA 831
661 GTAGAGAGTGGGTTCCAGAGACACCTGCTCCCGGTTGTGTCAGAGAGGCGCCGTA 720
832 GATGACACCTTGTGGCCACTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 891
721 GATGACACCTTGTGGCCACTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
892 GAGGAGCTTGTGCTAGTCCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 951
781 GAGGAGCTTGTGCTAGTCCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 840
952 GACCCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1011
841 GACCCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
1012 TCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
901 TCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
1072 GAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131
961 GAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1132 AAAGTGTGTTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1191
1021 AAAGTGTGTTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1080
1192 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
1081 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1252 CAGCAGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1284
1141 CAGCAGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1173

RESULT 13
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS AX127642
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source 1.1155
/organism="Homo sapiens"
/db.xref="taxon:9606"

BASE COUNT 185 a 349 c 382 g 239 t
ORIGIN
Query Match 73.4% Score 1155; DB 6; Length 1155;
Best Local Similarity 100.0% Pred. No. 3.6e-205;
Matches 1155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
130 ATGATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 189
1 ATGATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
190 CTGAACCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 249
61 CTGAACCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
250 CTTTGGCTGAGAGCTGAATCTCTTACAGCTGATGCTGACTGAGGAGGAGGAGGAGGAG 309

Db	121	CTTTGGCTAGAGCTGAATCTCTTCCAGCGTGTACTAGACGGGGACACGCG	180
Qy	310	CGGAGCTGGTGGGTGCGAGGAGCTGGGCGCTGGGAGCGCTGGTGTATGCTTGA	369
Db	181	CGGAGCTGGTGGGTGCGAGGAGCTGGGCGCTGGGAGCGCTGGTGTATGCTTGA	240
Qy	370	GACGGCTGATGACACAGTGTGTAAACGGGCTCATGTGACGGCGCTGTACTGGGAGACGCG	429
Db	241	GACGGCTGATGACACAGTGTGTAAACGGGCTCATGTGACGGCGCTGTACTGGGAGACGCG	300
Qy	430	ATCCAGAGCCCGTGGTACCCTCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTG	489
Db	301	ATCCAGAGCCCGTGGTACCCTCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTG	360
Qy	490	AACCATATGCTGGCTATGAGCAGAGTACCAATGAAGACCTCTTACCAACCTGACGCTA	549
Db	361	AACCATATGCTGGCTATGAGCAGAGTACCAATGAAGACCTCTTACCAACCTGACGCTA	420
Qy	550	TTTGCTGCGCGCGGCTGCTGTACACCCATGAACCTGTGTCTGTCTGACACGGCTTG	609
Db	421	TTTGCTGCGCGCGGCTGCTGTACACCCATGAACCTGTGTCTGTCTGACACGGCTTG	480
Qy	610	CTGGGCTCTTCTCTGTGTGTACACCTGGGCTGGGGCTTATGTGTGTGTGTGTGTGT	669
Db	481	CTGGGCTCTTCTCTGTGTGTGTACACCTGGGCTGGGGCTTATGTGTGTGTGTGTGT	540
Qy	670	AGTAGAAGATATGCGGTGTGGGGGAGATGCGCTTACTCTGGGAGACCTTCTGGCTGTG	729
Db	541	AGTAGAAGATATGCGGTGTGGGGGAGATGCGCTTACTCTGGGAGACCTTCTGGCTGTG	600
Qy	730	GCAGCCCTGTGCGACCTACCGCGGCGCATGGCCCTACCTCCCTGTAGAGAACTGGGCTTC	789
Db	601	GCAGCCCTGTGCGACCTACCGCGGCGCATGGCCCTACCTCCCTGTAGAGAACTGGGCTTC	660
Qy	790	AAGACACTGCTCCCTCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	849
Db	661	AAGACACTGCTCCCTCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	720
Qy	850	CTGAGAGAGCAGTAGTCCCTTCACTGTGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGT	909
Db	721	CTGAGAGAGCAGTAGTCCCTTCACTGTGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGT	780
Qy	910	CTGGACACTGTGCTACCTGTGACCTGGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	969
Db	781	CTGGACACTGTGCTACCTGTGACCTGGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	840
Qy	970	GCACCTGGGCTCATGTCTGT	1029
Db	841	GCACCTGGGCTCATGTCTGT	900
Qy	1030	CGGCTCTTCCGATGGGATGAGAGGCGAGGCAATATGGAATATGAAATGCCCACTTGTGTA	1089
Db	901	CGGCTCTTCCGATGGGATGAGAGGCGAGGCAATATGGAATATGAAATGCCCACTTGTGTA	960
Qy	1090	TATGTGCGCGT	1149
Db	961	TATGTGCGCGT	1020
Qy	1150	GATGGGAAATTTGATGTTTACCGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1209
Db	1021	GATGGGAAATTTGATGTTTACCGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1080
Qy	1210	ATGTTGTACGCGT	1269
Db	1081	ATGTTGTACGCGT	1140
Qy	1270	GAGAGCCCTTATGA 1284	
Db	1141	GAGAGCCCTTATGA 1155	

[illegible]


```
* 2878 4283: contig of 1406 bp in length
* 4284 4383: gap of 100 bp
* 4384 5447: contig of 1064 bp in length
* 5448 5547: gap of 100 bp
* 5548 6776: contig of 1229 bp in length
* 6777 6876: gap of 100 bp
* 6877 8215: contig of 1339 bp in length
* 8216 8315: gap of 100 bp
* 8316 9961: contig of 1646 bp in length
* 9962 10061: gap of 100 bp
* 10062 12224: contig of 2163 bp in length
* 12225 12324: gap of 100 bp
* 12325 13794: contig of 1470 bp in length
* 13795 13894: gap of 100 bp
* 13895 16857: contig of 2963 bp in length
* 16858 16957: gap of 100 bp
* 16958 18266: contig of 1309 bp in length
* 18267 18366: gap of 100 bp
* 18367 20190: contig of 1824 bp in length
* 20191 20290: gap of 100 bp
* 20291 21460: contig of 1170 bp in length
* 21461 21560: gap of 100 bp
* 21561 23407: contig of 1847 bp in length
* 23408 23507: gap of 100 bp
* 23508 25455: contig of 1948 bp in length
* 25456 25555: gap of 100 bp
* 25556 27331: contig of 1776 bp in length
* 27332 27431: gap of 100 bp
* 27432 30406: contig of 2875 bp in length
* 30407 30506: gap of 100 bp
* 30507 32805: contig of 2299 bp in length
* 32806 32905: gap of 100 bp
* 32906 35659: contig of 2754 bp in length
* 35660 35759: gap of 100 bp
* 35760 37974: contig of 2215 bp in length
* 37975 38074: gap of 100 bp
* 38075 40521: contig of 2447 bp in length
* 40522 40621: gap of 100 bp
* 40622 43219: contig of 2598 bp in length
* 43220 43319: gap of 100 bp
* 43320 46161: contig of 2842 bp in length
* 46162 46261: gap of 100 bp
* 46262 48463: contig of 2202 bp in length
* 48464 48563: gap of 100 bp
* 48564 51275: contig of 2712 bp in length
* 51276 51375: gap of 100 bp
* 51376 53522: contig of 2147 bp in length
* 53523 53622: gap of 100 bp
* 53623 56955: contig of 3333 bp in length
* 56956 57055: gap of 100 bp
* 57056 59555: contig of 2500 bp in length
* 59556 59655: gap of 100 bp
* 59656 62713: contig of 3058 bp in length
* 62714 62813: gap of 100 bp
* 62814 66519: contig of 3706 bp in length
* 66520 66619: gap of 100 bp
* 66620 68930: contig of 2311 bp in length
* 68931 69030: gap of 100 bp
* 69031 72724: contig of 3694 bp in length
* 72725 72824: gap of 100 bp
* 72825 76092: contig of 3268 bp in length
* 76093 76192: gap of 100 bp
* 76193 80143: contig of 3951 bp in length
* 80144 80243: gap of 100 bp
* 80244 84957: contig of 4714 bp in length
* 84958 85057: gap of 100 bp
* 85058 88424: contig of 3367 bp in length
* 88425 88524: gap of 100 bp
* 88525 93428: contig of 4904 bp in length
* 93429 93528: gap of 100 bp
* 93529 97177: contig of 3649 bp in length
* 97178 97277: gap of 100 bp
* 97278 102713: contig of 5436 bp in length
```

```
* 102714 102813: gap of 100 bp
* 102814 109141: contig of 6328 bp in length
* 109142 109241: gap of 100 bp
* 109242 114513: contig of 5272 bp in length
* 114514 114613: gap of 100 bp
* 114614 124525: contig of 9912 bp in length
* 124526 124625: gap of 100 bp
* 124626 132364: contig of 7739 bp in length
* 132365 132464: gap of 100 bp
* 132465 139488: contig of 7024 bp in length
* 139489 139588: gap of 100 bp
* 139589 148734: contig of 9146 bp in length
* 148735 148834: gap of 100 bp
* 148835 159465: contig of 10631 bp in length
* 159466 159565: gap of 100 bp
* 159566 172685: contig of 13120 bp in length
* 172686 172785: gap of 100 bp
* 172786 183443: contig of 10658 bp in length.
```

FEATURES

source

1. 183443

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone.lib="CITD Human BAC"

misc_feature 1. 1059

/note="assembly_fragment"

1160. 2777

/note="assembly_fragment"

2878. 4283

/note="assembly_fragment"

4384. 5447

/note="assembly_fragment"

5548. 6776

/note="assembly_fragment"

6877. 8215

/note="assembly_fragment"

8316. 9961

/note="assembly_fragment"

10062. 12224

/note="assembly_fragment"

12325. 13794

/note="assembly_fragment"

13895. 16857

/note="assembly_fragment"

16958. 18266

/note="assembly_fragment"

18367. 20190

/note="assembly_fragment"

20291. 21460

misc_feature

Query Match

Best Local Similarity 66.9%; Score 1052.8; DB 2; Length 183443;

Matches 1054; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 503 GCTATGACGAGCTGACCAATGAAGACCTCTGACACACTGACACGCTATTGCTGCGGCC 562
|||||
DB 134217 GCTATGACGAGCTGACCAATGAAGACCTCTGACACACTGACACGCTATTGCTGCGGCC 134158

QY 563 GGCTGCTGACCAATGAAGACCTCTGCTGACACGCTGCGGCTGCGGCTCTTCT 622
|||||
DB 134157 GGCTGCTGACCAATGAAGACCTCTGCTGACACGCTGCGGCTGCGGCTCTTCT 134098

QY 623 CTGTGCTGACCTGCGGCTGCGGCTCTTCTGCTGATGAGACCTAGAGAGTGAAGATATC 682
|||||
DB 134097 CTGTGCTGACCTGCGGCTGCGGCTCTTCTGCTGATGAGACCTAGAGAGTGAAGATATC 134038

QY 683 GCGCTTGGGGAGATGCGCTTCTGCTGAGCAGCTTCTGCGCTGCGGAGCCCTGCGCA 742
|||||
DB 134037 GCGCTTGGGGAGATGCGCTTCTGCTGAGCAGCTTCTGCGCTGCGGAGCCCTGCGCA 133978

QY 743 CTTACCGCGGCGACTGGGCTTACCTCTGCTGAGAGAGTGGGTTCCAGACACTGCGCT 802
```

```

Db 133977 CTTACCGCGGCGAGCTGGCCCTACCTCCCTGTGTGAGAAAGATGGGTTCCAGAGACACTTCCCT 133918
OY 803 CCCCCGTTGTGTCCAGAGAGGCCCCGGTGTAGATGACACACTTGTGCACTGTGAGAGAGCCAG 862
Db 133917 CCCCCGTTGTGTGTCCAGAGAGGCCCCGGTGTAGATGACACACTTGTGCACTGTGAGAGAGCCAG 133858
OY 863 TGGCCTCTCTACTGAGACAGTGGTGGCCGAGAGAGACTTGTGTCTGTGCACTGTGCACTGTGCA 922
Db 133857 TGGCCTCTCTACTGAGACAGTGGTGGCCGAGAGAGACTTGTGTCTGTGCACTGTGCACTGTGCA 133798
OY 923 ACTCGACCTGGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 982
Db 133797 ACTCGACCTGGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133738
OY 983 TGCATCTGTTTACGTGTGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1042
Db 133737 TGCATCTGTTTACGTGTGGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133678
OY 1043 CCATGGAGAAAGGCGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1102
Db 133677 CCATGGAGAAAGGCGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133618
OY 1103 TCGCCTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1162
Db 133617 TCGCCTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133558
OY 1163 TGGTTAGCAGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1222
Db 133557 TGGTTAGCAGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133498
OY 1223 GCGTGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1282
Db 133497 GCGTGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133438
OY 1283 GACCCCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1342
Db 133437 GACCCCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133378
OY 1343 GCTGCAGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1402
Db 133377 GCTGCAGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133318
OY 1403 GAAGGTGAGGCTATGCTTTGGGGGAGACAGCCAGATGAAGTCTGTGGTCCAGAGCCCA 1462
Db 133317 GAAGGTGAGGCTATGCTTTGGGGGAGACAGCCAGATGAAGTCTGTGGTCCAGAGCCCA 133258
OY 1463 GCTGTGCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1522
Db 133257 GCTGTGCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133198
OY 1523 CGAACCAAAATCAATAAAGTGACATTCCAGCGTG 1558
Db 133197 CGAACCAAAATCAATAAAGTGACATTCCAGCGTG 133162

```

Search completed: May 17, 2003, 13:39:06
 Job time : 4577 secs

Mon May 19 10:50:38 2003

us-09-937-060a-19.rni

Page 2

APPLICANT: Fuente, Juan L.
 APPLICANT: Liarena, Francisco J.
 APPLICANT: Liras, Paloma
 TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
 TITLE OF INVENTION: LATE GENES
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: John W. Wallen III
 STREET: P.O. Box 2000
 CITY: Rahway
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07065
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/04801
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Wallen III John W.
 REGISTRATION NUMBER: 5,403
 REFERENCE/DOCKET NUMBER: 19179
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 594-3505
 TELEFAX: (908) 594-4720
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 900 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 PCT-US95-04801-3

Query Match	3.0%	Score 47.8;	DB 5;	Length 900;
Best Local Similarity	51.1%;	Pred. No. 0.023;		
Matches 112; Conservative	0;	Mismatches 107;	Indels 0;	Gaps 0;

[illegible]

RESULT 3
 US-08-209-747-1
 Sequence 1, Application US/08209747
 Patent No. 5733771
 GENERAL INFORMATION:
 APPLICANT: Lewis, Randolph V.
 APPLICANT: Colgin, Mark
 TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
 TITLE OF INVENTION: Silk Proteins
 NUMBER OF SEQUENCES: 56
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Birch, Stewart, Kolasch & Birch
 STREET: P.O. Box 747
 CITY: Falls Church

```

STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"

```

Query Match	3.0%;	Score 47.2;	DB 1;	Length 2793;
Best Local Similarity	47.1%;	Pred. No. 0.044;		
Matches 145; Conservative	0;	Mismatches 163;	Indels 0;	Gaps 0;

QY	105	GCAGCGGAGAGCCCGCGGCTCAGGTTTATGATCTACAGGGGGCCCCCGGGCGGTGCTCC	164
Db	1331	GCAGGAGCCCGTGGGGGTGAGAGACTGAGGCTACGGTGGTCAAGTGGCTACGGTGCC	1390
QY	165	GGGGCCCTGGCGCGTGGTGGTGGTGTGAGAACCCGCGCGCGGCAAGGCGCTTCGA	224
Db	1391	GGTGGAGAGCTGGTGGCGGCTGCTGCTGCTGAGAGCAGAGCTGAGAGCGCTGTGGTTAC	1450
QY	225	GCCTTTCCGGAGTACAGCGAGCGCCCTTTGGGCTGAGGCTGAAATCTCCCTCACGCTAT	284
Db	1451	GGTAAAGTGGTGGTGGTGGAGCTGGAGCTGCTGAGAGCGCAGAGAGCTGAGAGCTACGCT	1510
QY	285	GCTCACATAGCGCGGGAACACAGCGCGGGAGCTGGTGGTCCGAGAGAGCTGGCGCCGTG	344
Db	1511	GGTCAAGTGATACGGGTGCGCGAGAGAGAGCTGGTGGTGGAGGCTGGTGCACAGAGA	1570
QY	345	GCAGCGCTTGGTGTCTATGCTCTCTGAGACGGGCTGATGCACAGAGTGGTGACAGCGCTAT	404
Db	1571	GGCGGAGCGCGCTGGTGCTTACGGTAAAGTGGTGGTGGTGGAGCTGGTGGCGCTCTGGG	1630
QY	405	GGAGCGGC	412
Db	1631	GCAGGTGC	1638

RESULT 4
 US-08-458-298-1
 ; Sequence 1, Application US/08458298
 ; Patent No. 5756677
 ; GENERAL INFORMATION:
 ; APPLICANT: Lewis, Randolph V.
 ; APPLICANT: Colgin, Mark

GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Armentrout, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT APPLICATION NUMBER: US/09/096,867
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 60/049,428
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 28804
TYPE: DNA
ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Query Match
Best Local Similarity 49.5%; Pred. No. 0.88;
Matches 110; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

OY 164 GCGGCGCCCTGCGCCCTGCTGCTGCTGAACCGCGCGCGGCAAGGCGACCTTGC 223
DB 28003 GCGGCGCCCTGCGCGCGCTGCTGCTGCGCGCGCGGCGGCGGCTGCGCGCGGCTGCA 28062
OY 224 AGCTCTTCGAGAGTACGACGACCGCTTTTGGCTGAGCTGAATCTCTTCACCGCTCA 283
DB 28063 CTTTCGCGCGGCTGATCTGCTGCGCGCGCGCGGCGGCGGCGGCTGCGCGCGGCTG 28122
OY 284 TGTCTACTGAGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 343
DB 28123 CATCTTCGCGCGGCTGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 28182
OY 344 GGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
DB 28183 GGGGCGCGAGTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 28224

RESULT 11

US-08-804-227C-1

Sequence 1, Application US/08804227C

Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Kinstross, Stuart A.

APPLICANT: Rostek, Paul R., Jr.

APPLICANT: Sulton, Kimberly L.

TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: THOMAS G. PLANT 1501

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCI(DOS) Text only

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,227C

FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.

REGISTRATION NUMBER: 35,784

REFERENCE/DOCKET NUMBER: X-8231

TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43280 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 816..14234
FEATURE:
NAME/KEY: CDS
LOCATION: 14351..19945
FEATURE:
NAME/KEY: CDS
LOCATION: 20010..31199
FEATURE:
NAME/KEY: CDS
LOCATION: 31232..36067
FEATURE:
NAME/KEY: CDS
LOCATION: 36249..41774
US-08-804-227C-1

Query Match
Best Local Similarity 48.0%; Pred. No. 0.98;
Matches 122; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

OY 145 GCGCCCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204
DB 12177 GGGGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12236
OY 205 GCGAAGGCGAAGCGCTTGCAGCTTGCAGCTTGCAGCTTGCAGCTTGCAGCTTGCAGCTTGC 264
DB 12237 GCGGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12296
OY 265 GAATCTCTTACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 324
DB 12297 CCGCGCTCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12356
OY 325 TCGGAGGAGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 384
DB 12357 GGGGATCAGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 12416
OY 385 GAGGTGTGAGCGG 398
DB 12417 GCGCTGCGCGCGG 12430

RESULT 12

US-08-949-155-49/C

Sequence 49, Application US/08949155

Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A

APPLICANT: Bazer, Fuller W

TITLE OF INVENTION: Compositions and Methods for the

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE AND DURKEE

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

OY 180 GCTGCTGCTGCTGAAACCCGCGGCGGCAAGCCGCTTGACGCTTTCGGAGTCA 239
 DB 226 GCTGCTGCTGCTGAAACCCGCGGCGGCGGCAAGCCGCTTGACGCTTTCGGAGTCA 285
 OY 240 CCTGACACCCCTTTTGGCTGAGGCTGAATTCCTTACGCTGATGCTCACTGAGCGCG 299
 DB 286 CGTGACACCCCTTTTGGCTGAGGCTGAATTCCTTACGCTGATGCTCACTGAGCGCG 345
 OY 300 GAACACACCCGCGGAGCTGCTGCGCTGAGAGAGCTGCGCGCTGGAGACCTTGAGTGT 359
 DB 346 GAACACACCCGCGGAGCTGCTGCGCTGAGAGAGCTGCGCGCTGGAGACCTTGAGTGT 405
 OY 360 CATGCTGAGAGAGGCTGATGACACAGAGTGTGAACGGGCTCAGAGACGCGCTGACTG 419
 DB 406 CATGCTGAGAGAGGCTGATGACACAGAGTGTGAACGGGCTCAGAGACGCGCTGACTG 465
 OY 420 GGACACCGGCTATCCAGAACCCCTGTGTAGCTCCAGAGGCTCTGGCAAGCGCTGAC 479
 DB 466 GGACACCGGCTATCCAGAACCCCTGTGTAGCTCCAGAGGCTCTGGCAAGCGCTGAC 525
 OY 480 AGCTTCTGAAACCATTTATGCTGCTATGAGCAGAGTGTGACCAATGAAGACCTTGACCA 539
 DB 526 AGCTTCTGAAACCATTTATGCTGCTATGAGCAGAGTGTGACCAATGAAGACCTTGACCA 585
 OY 540 CTGACACCTTATGCTGCTGCGCGGCTGCTGCTGACACCAATGAAGACCTTGCTGACAC 599
 DB 586 CTGACACCTTATGCTGCTGCGCGGCTGCTGCTGACACCAATGAAGACCTTGCTGACAC 645
 OY 600 GCGTTCGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 DB 646 GCGTTCGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
 OY 660 GGACCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGACCTT 719
 DB 706 GGACCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGACCTT 765
 OY 720 CCGTGGCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGAG 779
 DB 766 CCGTGGCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGAG 825
 OY 780 AGTGGGTTTCAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
 DB 826 AGTGGGTTTCAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
 OY 840 CCGTGGCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGAG 899
 DB 886 CCGTGGCTGAGAGAGTGAAGATATCGGCGCTGCTGGGAGAGATCGCTTCACTGCTGAGAG 945
 OY 900 TGTGCTAGTCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
 DB 946 TGTGCTAGTCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
 OY 960 GGGCGGCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
 DB 1006 GGGCGGCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
 OY 1020 CATGCTGCTGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
 DB 1066 CATGCTGCTGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
 OY 1080 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
 DB 1126 CTACTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1185
 OY 1140 GTTTCAGTATGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
 DB 1186 GTTTCAGTATGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1245
 OY 1200 CTACTTGTATGATGCTGAGCGGTTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
 DB 1246 CTACTTGTATGATGCTGAGCGGTTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1305

OY 1260 GCCACGCGCAGAGAGAGCGCTTATGACCCCTGAGCGGCTGAGCTTACTGCTACTGCTC 1319
 DB 1306 GCCACGCGCAGAGAGAGCGCTTATGACCCCTGAGCGGCTGAGCTTACTGCTACTGCTC 1365
 OY 1320 AGGACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
 DB 1366 AGGACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1425
 OY 1380 GAGACTCTCTGAGAGAGGCTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1439
 DB 1426 GAGACTCTCTGAGAGAGGCTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1485
 OY 1440 TGAAGTCTGAG 1499
 DB 1486 TGAAGTCTGAG 1545
 OY 1500 TTTGTTCTGAG 1553
 DB 1546 TTTGTTCTGAG 1599

RESULT 2
 US-09-970-516-1
 : Sequence 1, Application US/09970516
 : Patent No. US20020099029A1
 : GENERAL INFORMATION:
 : APPLICANT: No. US20020099029A1a1a1s AG
 : TITLE OF INVENTION: Induction of blood vessel formation through administration of
 : FILE REFERENCE: 4-31617
 : CURRENT FILING DATE: 2001-10-04
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 1
 : LENGTH: 1155
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)..(1155)
 : OTHER INFORMATION:
 : US-09-970-516-1

Query Match 73.4%; Score 1155; DB 10; Length 1155;
 Best Local Similarity 100.0%; Pred. No. 2,2e-313;
 Matches 1155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 ATGATTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
 DB 130 ATGATTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190
 OY 190 CTGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 249
 DB 61 CTGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 OY 250 CTTTGGCTGAGGCTGAATTCCTTACGCTGATGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 309
 DB 121 CTTTGGCTGAGGCTGAATTCCTTACGCTGATGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 180
 OY 310 CGGAGAGTGTGCGGCTGAGAGAGTGTGCGGCGGCTGAGAGAGTGTGAGAGTGTGAGAG 369
 DB 181 CGGAGAGTGTGCGGCTGAGAGAGTGTGCGGCGGCTGAGAGAGTGTGAGAGTGTGAGAG 240
 OY 370 GAGGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 429
 DB 241 GAGGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 300
 OY 430 ATCCAG 489
 DB 301 ATCCAG 360
 OY 490 AACCATTTATGCTGCTATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 549

```

Db 361 AACCATTTATGCTGAGGATGAGGAGTACCAATGAAGACCTCTGACCACTGACAGCTA 420
QY 550 TTGCTGCGCGCGGCTGCTGACACCAATGAAGACCTCTCTGACACGCGCTTCGGGG 609
Db 421 TTGCTGCGCGCGGCTGCTGACACCAATGAAGACCTCTCTGACACGCGCTTCGGGG 480
QY 610 CTGCGCCCTCTCTGCTGCTGACGCTGCGCTGCGGCTCATTTGCTGATGTGACCTAGAG 669
Db 481 CTGCGCCCTCTCTGCTGCTGACGCTGCGCTGCGGCTCATTTGCTGATGTGACCTAGAG 540
QY 670 AGTGAAGATATGCGCGCTGCGGAGATGCGCTTCTACTCTGCGGACCTTCTGCGCTG 729
Db 541 AGTGAAGATATGCGCGCTGCGGAGATGCGCTTCTACTCTGCGGACCTTCTGCGCTG 600
QY 730 CGAGCGCTGCGGACCTGACGCGCGCGGACCTGCGCTACTCTCTGTAAGAGATGCGCTTC 789
Db 601 CGAGCGCTGCGGACCTGACGCGCGCGGACCTGCGCTACTCTCTGTAAGAGATGCGCTTC 660
QY 790 AAGACACCTGCTCCCGCTGCTGCTGACGAGGCGCGGCTGATGACACCTGCTGCGCA 849
Db 661 AAGACACCTGCTCCCGCTGCTGCTGACGAGGCGCGGCTGATGACACCTGCTGCGCA 720
QY 850 CTGAGAGAGACCTGACGCGCTCTGCTACTGAGACAGTGTGCCGAGAGACCTTGTGCTAGTC 909
Db 721 CTGAGAGAGACCTGACGCGCTCTGCTACTGAGACAGTGTGCCGAGAGACCTTGTGCTAGTC 780
QY 910 CTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
Db 781 CTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 970 CGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
Db 841 CGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1030 CGCCTCTCTCTGCGCATGAGAGAGGCGAGCATATGAGATGATGATGATGATGATGATG 1089
Db 901 CGCCTCTCTCTGCGCATGAGAGAGGCGAGCATATGAGATGATGATGATGATGATGATG 960
QY 1090 TATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
Db 961 TATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1150 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1209
Db 1021 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1210 ATGCTCAGCGGTTGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1269
Db 1081 ATGCTCAGCGGTTGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1140
QY 1270 GAAGAGCCCTTATGA 1284
Db 1141 GAAGAGCCCTTATGA 1155

```

RESULT 3 US-09-784-810A-3

```

; Sequence 3, Application US/09784810A
; Patent No. US2002082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 3
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-784-810A-3

```

```

Query Match          54.2% Score 853.2; DB 10: Length 1759;
Best local similarity 73.9% Pred. No. 6.3e+229;
Matches 1158; Conservative 0; Mismatches 388; Indels 22; Gaps 5;

```

```

QY 3 CCCACAGCCGCCCTGCGAGCGCCCGCTGCGGACGACCCGATGAAGAGCTGAAGGACGAG 62
Db 196 CCCACAGCCGCCCTGCGAGCGCCCGCTGCGGACGACCCGATGAAGAGCTGAAGGACGAG 255
QY 63 CCGCCGCCACCGGCGACCGCGCC---ACAGCCCGACGAGACCCCGCTGCGACCGCGGACCGCC 118
Db 256 CCGCCGCCCTGCTGAGCAGCGCGCGGCGGACGACCGGCTGCGCCCGCTGCTGCGAGCGGACCGCC 315
QY 119 GGTGCGAGGTTATGATTCAGCGGCGGCGCCCGCGGCGCTGCTCCCGCGCGCTGCGCGG 178
Db 316 GGTGCGAGGTTATGATTCAGCGGCGGCGCCCGCGGCGCTGCTCCCGCGCGCTGCGCGG 375
QY 179 TGTGTGCTGCTGAACCCGCGCGCGGCGGACGAGGCGCTTGCAGCTCTTCGCGAGTC 238
Db 376 TGTGTGCTGCTGAACCCGCGCGCGGCGGACGAGGCGCTTGCAGCTCTTCGCGAGTC 435
QY 239 AGCTGACGCGCTTTTGTGCTGAGGCTGAATCTCTTACGCTGATGCTGCTGACGCGC 298
Db 436 GTGTGCGAGCGCTTCTGCGAGGCGGAGAGATTAACCTTAACTGAATACCTGCGAGCGCA 495
QY 299 GGAACACGCGCGGCGAGCTGCTGCGGAGAGCTGCGGCGCGCTGCGGAGCTGCTGCTG 358
Db 496 AGAACCTATGCGAGGAGCTGCTGCTGCGAGAGAGTGTGGCTGCTGCGGAGCGCTGCGAG 555
QY 359 TCATGTCTGAGAGCGGCTGATGACAGAGGTGTGTAACGCGCTCATGAGAGCGGCTGACT 418
Db 556 TCATGTCTGAGAGCGGCTGATGACAGAGGTGTGTAACGCGCTCATGAGAGCGGCTGACT 615
QY 419 GGGAGACCGCGCATTCAGAGAGCGCGCTGCTGAGCGCTCCGACAGCGCTGCTGCGAGCG 478
Db 616 GGGAGACCGCGCATTCAGAGAGCGCGCTGCTGAGCGCTCCGAGAGCGCTGCGAGAGCG 675
QY 479 CAGCTTCTTGAACCATATGATGCTGCTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 538
Db 676 CAGCTTCTTGAACCATATGATGCTGCTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 735
QY 539 ACTGACAGCTATTCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
Db 736 ACTGACAGCTATTCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
QY 599 CGGCTTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
Db 796 CGGCTTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
QY 659 TGGACCTAGAGAGTGAAGATATGCGGCTGCTGCGGAGATGCTGCTGCTGCTGCTGCTGCTG 718
Db 856 TGGACCTAGAGAGTGAAGATATGCGGCTGCTGCGGAGATGCTGCTGCTGCTGCTGCTGCTG 915
QY 719 TCTGTGCTGCTGCGAGCGCTGCGGACCTGACCGGCGGAGCTGCGCTGCTGCTGCTGCTG 778
Db 916 TCTGTGCTGCTGCGAGCGCTGCGGACCTGACCGGCGGAGCTGCGCTGCTGCTGCTGCTG 975
QY 779 GATGTGCTTCAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
Db 976 GATGTGCTTCAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032
QY 839 ACCTTGTGCGACCTGCGAGGAGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898
Db 1033 ACCTTGTGCGACCTGCGAGGAGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
QY 899 TTGTGCTAGTCTGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
Db 1093 TTGTGCTAGTCTGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152

```

Mon May 19 10:50:38 2003

us-09-937-060a-19.rnpb

Page 4

QY	959	TGGCCGCGTGCACATGGCGTCATGATCTGTTCACGTGGCGGCGAGTGTCTCGTG	1018
Db	1153	TGGCGCCGCTGTAGAGCTGTGGTTATGCAATCTGTTCACGTAGTGTGGGGGTGTCAAAGGG	1212
QY	1019	GCATGCTGCTGGCCTCTTCTGGCCATGAGAGGCGAGCGATATGAGATGAATGGC	1078
Db	1213	CTCGCGTGTGGCCCTCTTCTGGCCATGTGAGAAAGGCAAGCATATGAACTTGACTGTTC	1272
QY	1079	CCATCTTGGTATATCTGCCCGTGGTCCCTCTTCCGCTTGGAGGCCCAAGATGGGAAAGGTG	1138
Db	1273	CATACCTGGTTATGTGGCCCGTGGTTCCTTCCGCTGTGAGGCCAGAGCCAAAGGGGCG	1332
QY	1139	TGTTTCAGTGTGATGGGATATTGATGGTTAGCGAGAGCCGTGCAGAGCCAGGTGCACCCAA	1198
Db	1333	TGTTTCTGTGATGAGAAAGCTGATGGTATGTGAACCTGTGAGGGCCCAATGTCACCCAA	1392
QY	1199	ACTACTCTGGATGTGTGACGGTTCGCTGGAGAGCCCGCCGACGGAGGGAAGCCGACACAA	1258
Db	1393	ACTACCTTTGGATGGTCTGTGTGGAGAGAGAAATGCCCCATCCGGCCGGACCTCCGGCGGG	1452
QY	1259	TGGCACCAGGCAAGAGGCTCT--TATGACCCCTGGAGCGCGGCTGTGCCTTAAGTGTAC	1315
Db	1453	GGGCACCCTCAAGAAAGCATATACCTGTGACCTTGTCTACTCTGTGTAGGCTGACATGG	1512
QY	1316	TTCGAGAGCCCTTCCTCCCTCCCTAGGCGTCGAGGCGCTGTCCACAGCTCTGTGGGGGT	1375
Db	1513	GACCCCTCCCCCAGCCACCCACCTCTGTGATGTGGAGGTATTCTTAAAGTTCCTATGCMACT	1572
QY	1376	GGAGGAGACTCTCTGTGAGAGAGGTGAGAGGTGAGAGGTATGCTTTGGGGGAGACGGCC	1435
Db	1573	GGTGGGAGACCCCTGTGAGAGAAAGCT--AGAAGGTGGGCTGTGACTTGTGAAGAAAGAGGCT	1630
QY	1436	AGATGAATCTCTGTGTAGAGGCCAGCT-----GGCTGGGCGCCAGGTGGCTAT	1485
Db	1631	TTTACCTTCAGTTAGATGATGATCATCCCAATGAGAGCCCTGTGTGGTGGACCAAGTTGCATTT	1690
QY	1486	GTTAGGCGCTTTAGTTTCTGTGAGAGCCGACCCACAGAAACAAATCAAAATTAAGTGA	1545
Db	1691	AGAAAGCATTTCCCACTTCTTTTAGAGACCTTCCCTGGAAACCAATTCAAATTAAGAGA	1750
QY	1546	CATTCCCA 1553	
Db	1751	CTTTTCCA 1758	

```

RESULT 4
US-09-970-516-5
Sequence 5, Application US/0970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1ar1s AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-11617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:
US-09-970-516-5

```

Query Match	48.8%	Score 766	DB 10	Length 1149
Best Local Similarity	79.8%	Pred. No. 3.9e-205		
Matches 919	Conservative	0	Mismatches 230	Indels 3
				Gaps 1

QY	130	ATGATGACACAGCGGCGGCGCCCGCGGCGGTGCTGCCGCGCCCTGCGCGTGTGTGCTG	189
Db	1	ATGCAGACCGATGGAATGGCCCGCGAGGAGCTGTCCACGCGCATCAAGTGTGTGTG	60
QY	190	CTCAACCCCGCGCGGCGGACAGGCGCATGGCTTGGACAGCTCTTCCGAGATCAGTGCAGGCC	249
Db	61	CTGAACCCCGCAGGGTGGCAAGGAGTAAGCTGTGCAGCTCTTCCAGAGCCGTGTGCAGGCC	120
QY	250	CTTTTGGCTGGAGGCGCAATCTCTCTTCCACGCTGATGTGTACTGAGCGCGGACACACGG	309
Db	121	TTCTCTGGAGAGGCGAGAGATTAACCTTTAACTGTATATCTACCGAAGCAAGAACATATCC	180
QY	310	CGGAGACCTGTGGCGCTCGAGAGAGCTGGCGCGCTGGAGAGCTGTGTGTATGTGTGA	369
Db	181	AGGAGACTGTGTGTGTGCAGAGCAATTTGGTGCATGTGGAGCGCTTGGGAGATATGTCCGT	240
QY	370	GACGGCTGTATGCACAGAGTGTGTGAAGCGGCTCATGAGCGCGCTGACTGGAGACGCC	429
Db	241	GATGGCTGTATGTGCATGAGGTGTGTGAATGTGGATTAATGAAGACGGCCCGATGGAGACTGTCC	300
QY	430	ATTCAGAAAGCCCTGTGTGATACCTCCAGAGGCTCTGGCAAGCGCGCTGGACGCTTCTTG	489
Db	301	ATTCAGAAAGCCCTGTGTGATACCTCCCTGGAGGCTCTCGGCAATGGCGCTGGACGCTTCTTG	360
QY	490	AACCATTAATGCTGTGCTATGTAGCAGAGTGCACCAATGAAAGACCTCTGACCAACTGTCAAGCTA	549
Db	361	AACCATTAATGCTGTGGGTATGAGGCGAGGAGCTGACTAATGAAAGACCTGTCAACTGTCAAGCTG	420
QY	550	TTGCTGTGCGCGCGGCTGAGCTGATACCACATGAACCTGTGTCTGTGCACACGCGCTTGGCGG	609
Db	421	CTGTGTGTGCGCGCGCGCGCTGTACCCATGAACCTGTGTCTGTGTGCACACTGTGTGTGG	480
QY	610	CTGCGCTCTCTCTGTGTGTGCACAGCTGTGGCTGTGGGCTCATGTGTGTGTGTGACACCTAAG	669
Db	481	CTTCGCGCTCTATCTGTGTGTGCAGTGTCTCTGGGCTTTGTGTGTGACGTGTGACCTGTAG	540
QY	670	ACGTGAAGATATCGGCTGTGGGGGAGATGTGGCTTCACTGTGGGACACTTCTGTGGCTGTG	729
Db	541	ACGTGAAGATATACAGGCTTGTGGGGAGATTTCTTTCAAGTGTGGACCTTCTTTCGCTTA	600
QY	730	GGAGCGCTGTGGCACTTACGCGCGCGCGACAGTGGCTTACCTCTCCGTGAAGAAAGTGGGTGC	789
Db	601	GCAGCTGTGTGCGATGTATACCAAGGCTACAGGCTACCTCTCTCTGTGAAGAACTGTGGCTCT	660
QY	790	AAGACACTGCCCTCCCGCTTGTGTGTGCAGCAGGCGCGGTAGATGTGCACACTTGTGCCA	849
Db	661	AAGAGACCGCGCTCTAC--ACGTGTGTGCAGAAAGGCGCGCGGTGCACACACTTGTTCCT	717
QY	850	CTTGAAGAGGACAGTGCCTCTCTCACTGTGACAGTGTGTCCCGCAGAGACTTGTGTACTC	909
Db	718	CTTGAAGAGGACAGTGCCTCTCTCTCACTGTGACTGTGTACAGAACAGGACTTGTGTCTGTG	777
QY	910	CTTGAAGACTGTGCACCTGTGCACCTGTGGCACTGTGAAGATTTGCTGTACCCATGTGGCGCTCT	969
Db	778	CTTGTGTGTGTGTACACACCACTGTGACCTGTGAGCTGTTTTGCAGACCCATGTGGCGCTCT	837
QY	970	GGAGCTGGCTCATGTCTGTCTACGTGTGGCGGCGGAGTGTCTCTGCACTGATGTGTG	1029
Db	838	GAGCGTGTGTATGTCACTCTCTCTACGTGTGTGCTGGGGGTGTCAAGGCGTGTGCGCTGTG	897
QY	1030	CGCTCTTCTGTGGCATGTGAAGAGGCGCAGGCAATGAGATGTGAATCCCTCACTTGTGTA	1089
Db	898	CGCTCTTCTGTGGCATGTGAAGAGGCGCAGGCAATGTGAATCTGTGATGTGTCTATCTGTGT	957
QY	1090	TATGTGTGCGGTGTGTGCGCTTGTGGCTTGTGAAGCCCAAGATGTGGAAAGTGTGTGTGAGTG	1149
Db	958	CATGTGTGCGGTGTGTGCTTGTGCGCTGTGAAGCCCAAGAGCCAGAGGCGGTGTGTGTGTG	1017
QY	1150	GATGGGAAATTTGATGTGTAGCGAGGCGCGTGTGAAGGCGCAGGTGTGCACCCAACTACTTGTGC	1209
Db	1018	GATGAGAGGAGGCTGTGTGTATGTAGTGTGTGACGGGCGCAAGTGCACCCAACTACTTGTGC	1077
QY	1210	ATGTGTACAGGTTGTGTGTGAGCCCGCGCCACATGTGGAAGCCCAAGCAAGATGTGCACCGCA	1269

QY	1165	GTTAGCAGAGCCGTGACAGGGCCAGAGTGCACCCAAATACTTCTGGATGGGTACAGGGTTGC	1224
Db	394	GTTACCGAGGCGCCGTGACAGGGCCAGAGTGCACCCAAATACTTCTGGATGGGTACAGTGTGGC	335
QY	1225	GTGAGAGCCGCCGCCACAGTGGAAAGCCCCAGACAGATGCCACGCCCAGAAAGAGCCCTTATGA	1284
Db	334	GTGAGAGCCGCCGCCACAGTGGAAAGCCCCAGACAGATGCCACGCCCAGAAAGAGCCCTTATGA	275
QY	1285	CCCCTGGGCGGCGCTGTGCTTATGTGTCTAATTGACAGAACCCCTTCCTCCTTAGAGGC	1344
Db	274	CCCCTGGGCGGCGGCTGTGCTTATGTGTCTAATTGACAGAACCCCTTCCTCCTTAGAGGC	215
QY	1345	TGCAGGGCGCTGTGCACAGGTCCTGTGGGGGTGGAGAGAGCTCCTCTGGAGAGGGGTGAGA	1404
Db	214	TGCAGGGCGCTGTGCACAGGTCCTGTGTGGGGGTGGAGAGAGCTCCTCTGGAGAGGGGTGAGA	155
QY	1405	AGGTGAGAGGCTATGCTTTTGGGGGGAGACAGGCGCAAAATGAAGTCTCTGGGTACAGAGCCACAG	1464
Db	154	AGGTGAGAGGCTATGCTTTTGGGGGGAGACAGGCGCAAAATGAAGTCTCTGGGTACAGAGCCACAG	95
QY	1465	TGGCTGGGCGCCAGCTGCTTATGTAAAGCCCTTAGTTTGTCTGAGACCCCAACCCCAAC	1524

```

RESULT 6
US-09-783-590-9248
: Sequence 9248, Application US/097835590
: Patent No. US20020110850A1
: GENERAL INFORMATION:
: APPLICANT: Dillon, Patrick J.
: APPLICANT: Haseltine, William A.
: APPLICANT: Li, Haodong
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
: FILE REFERENCE: PO-16.2c1
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 08/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/346,731
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9248
: LENGTH: 480
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (68)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (193)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (220)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (221)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (236)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (269)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (297)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (302)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (322)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (337)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (350)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (356)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (367)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (382)

```

```

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (385)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (391)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (400)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

```

```

Query Match 18.5%; Score 291; DB 10; Length 480;
Best Local Similarity 86.4%; Pred. No. 9,2e-72;
Matches 400; Conservative 0; Mismatches 55; Indels 8; Gaps 8;

```

```

OY 266 AATCTCTCTTCAAGCTGATGCTGAGCGGCGGAAACGCGGCGGAGCTGCTGCGCT 325
DB 10 AATCTCTCTTCAAGCTGATGCTGAGCGGCGGAAACGCGGCGGAGCTGCTGCGCT 69
OY 326 CGAGAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 385
DB 70 CGAGAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 129
OY 386 AG-CTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 442
DB 130 AGAGTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
OY 443 TGTGTACCTCTTCAAGCTGATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
DB 190 TGTGTACCTCTTCAAGCTGATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 249
OY 503 GCTATG-AGCAGGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 558
DB 250 GCTATGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 309
OY 560 GCGGAGCTGCTGACCAT-CAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
DB 310 GCGGAGCTGCTGACCATCAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
OY 619 TTTCTGTG-CTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 677
DB 370 TTTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 429
OY 678 GTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 430 GTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 472

```

```

RESULT 7
US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. US20020099029A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029A1artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1857

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)-(1857)
OTHER INFORMATION:
US-09-970-516-3

```

```

Query Match 16.7%; Score 262.8; DB 10; Length 1857;
Best Local Similarity 63.7%; Pred. No. 9,4e-64;
Matches 399; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

```

```

OY 147 CCCCCTGGGCGGCTGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 206
DB 408 CACCCCTGACCTGCTACCTGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 467
OY 207 CAAGGCGAAGGCGCTTGCAGCTCTTCCGAGTCACTGACGCGGCGGCGGCGGCGGCGG 266
DB 468 TCGGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
OY 267 AATCTCTTCAAGCTGATGCTGAGCGGCGGAAACGCGGCGGAGCTGCTGCGCTG 326
DB 528 GCTGTCTTCAAGCTGATGCTGAGCGGCGGAAACGCGGCGGAGCTGCTGCGCTG 587
OY 327 GAGAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 386
DB 588 GCTGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 647
OY 387 GGTGTGAAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 446
DB 648 GGTGTGAAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 707
OY 447 TACCTCTTCAAGCTGATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 506
DB 708 CATCTCTTCAAGCTGATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 767
OY 507 TGAAGAGTCAACCAATGAGAGCTCTGACCAACTGACGCTATTTGCTGCTGCTG 566
DB 768 TGAAGAGTCAACCAATGAGAGCTCTGACCAACTGACGCTATTTGCTGCTGCTGCTG 827
OY 567 GCTGTACCATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
DB 828 TGCACCACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
OY 627 GCTACACCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
DB 888 CTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947
OY 687 TCTGGGAGAGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
DB 948 CTGGGAGAGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1007
OY 747 CCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772
DB 1008 CCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033

```

```

RESULT 8
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegell, Sarah
; APPLICANT: Kohama, Taketumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; NUMBER OF SEQ ID NOS: 15
; PRIOR FILING DATE: 2000-04-03
; SOFTWARE: Patentin Ver. 2.0

```

```

: SEQ ID NO 13
: LENGTH: 2380
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (7)..(1860)
: PUBLICATION INFORMATION:
: TITLE: Molecular cloning and functional characterization of a
: TITLE: novel mammalian sphingosine kinase type 2 isoform
: JOURNAL: J Biol. Chem.
: VOLUME: 275
: ISSUE: 26
: PAGES: 19513-19520
: DATE: 2000-06-30
: DATABASE ACCESSION NUMBER: AF245447
: DATABASE ENTRY DATE: 2000-06-27
: US-09-817-676a-13

```

```

Query Match          16.7%; Score 262.8; DB 10; Length 2380;
Best Local Similarity 63.7%; Pred. No. 9.9e-64;
Matches 399; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

```

```

Oy 147 CCCCCGGGCGTCTCCGCGGCCCTGCGCGTGTGTCTGTGAACCCGCGCGCG 206
Db 414 CACCCCTGACCTGTACCTCGCGGCCCGCGTGTCTGTGAATCCCTTTGGGGG 473

Oy 207 CAAGGGCAAGGCGCTTCCGAGTCACCTGACGCCCTTTGGTGGAGGCTGA 266
Db 474 TCGGGGCGTGGCTCGGAGTGTGAAGACCCGCTTCCAGATCTCTGAAGCTGG 533

Oy 267 AATCTCTTCAACGATGCTCTACTGACGCGGGAACCAACGCGCGAGTGTGCGGTC 326
Db 534 GCTGTCCTTCAACCTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 593

Oy 327 GGAGAGCTGGGCGGCTGGAGAGCTCTGTGTGTCTGTGTGTGTGTGTGTGTGT 386
Db 594 GCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 653

Oy 387 GGTGTGTGAAGCGGCTGATGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT 446
Db 654 GGTGTGTGAAGCGGCTGATGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT 713

Oy 447 TAGCTTCCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 506
Db 714 CATCTTCCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 773

Oy 507 TGAGAGCTCAACCAATGAAGACCTCTGACCACTGACAGCTATTGTGTGTGTGT 566
Db 774 TGAGAGCTCAACCAATGAAGACCTCTGACCACTGACAGCTATTGTGTGTGTGT 833

Oy 567 GCTGTACCAATGAAGACCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 626
Db 834 TGAGAGCTCAACCAATGAAGACCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 893

Oy 627 GCTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 686
Db 894 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 953

Oy 687 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 746
Db 954 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1013

Oy 747 CCGGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 772
Db 1014 CCGGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1039

```

```

RESULT 9
US-09-817-676a-11
: Sequence 11, Application US/09817676A
: Patent No. US20020042101A1
: GENERAL INFORMATION:

```

```

: APPLICANT: Spiegel, Sarah
: APPLICANT: Kohama, Takatomi
: TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
: TITLE OF INVENTION: Expression and Methods of Use Thereof
: FILE REFERENCE: 00170/HG
: CURRENT APPLICATION NUMBER: US/09/817, 676A
: PRIOR FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/194, 318
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patentl Ver. 2.0
: SEQ ID NO 11
: LENGTH: 2698
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (387)..(2237)
: PUBLICATION INFORMATION:
: TITLE: Molecular cloning and functional characterization of a
: TITLE: novel mammalian sphingosine kinase type 2 isoform
: JOURNAL: J Biol. Chem.
: VOLUME: 275
: ISSUE: 26
: PAGES: 19513-19520
: DATE: 2000-06-30
: DATABASE ACCESSION NUMBER: AF245448
: DATABASE ENTRY DATE: 2000-06-27
: US-09-817-676a-11

```

```

Query Match          15.7%; Score 246.8; DB 10; Length 2698;
Best Local Similarity 62.1%; Pred. No. 3e-59;
Matches 389; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

```

```

Oy 147 CCCCCGGGCGTCTCCGCGGCCCTGCGCGTGTGTGTGTGTGTGTGTGTGTGTGT 206
Db 797 CACCCCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856

Oy 207 CAAGGGCAAGGCGCTTCCGAGTCACCTGACGCCCTTTGGTGGAGGCTGA 266
Db 857 GCGGGGCGTGGCTCGGAGTGTGAAGACCCGCTTCCAGATCTCTGAAGCTGG 916

Oy 267 AATCTCTTCAACGATGCTCTACTGACGCGGGAACCAACGCGCGAGTGTGCGGTC 326
Db 917 GCTGTCCTTCAACCTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 976

Oy 327 GGAGAGCTGGGCGGCTGGAGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
Db 977 GTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1036

Oy 387 GGTGTGTGAAGCGGCTGATGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT 446
Db 1037 GGTGTGTGAAGCGGCTGATGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGT 1096

Oy 447 TAGCTTCCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 506
Db 1097 TGTCTTCCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1156

Oy 507 TGAGAGCTCAACCAATGAAGACCTCTGACCACTGACAGCTATTGTGTGTGTGT 566
Db 1157 TGAGAGCTCAACCAATGAAGACCTCTGACCACTGACAGCTATTGTGTGTGTGT 1216

Oy 567 GCTGTACCAATGAAGACCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 626
Db 1217 CAGCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1276

Oy 627 GCTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 686
Db 1277 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1336

Oy 687 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 746
Db 1337 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1396

```

Mon May 19 10:50:38 2003

us-09-937-060a-19.inpb

Page 8

OY 747 CCGGAGCGAGCGGCTACGTCCTG 772
Db 1397 CCGTGGAGCGCTCTCTACCTCCCG 1422

RESULT 10
US-09-796-692-2905

Sequence 2905, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077, 001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-2905

Query Match 12.5% Score 197.4; DB 9; Length 199;
Best Local Similarity 99.5% Pred. No. 1.2e-45;
Matches 198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1043 CCATGAGAGAGCGGAGCATATGAGTATGATGCCCCCTACTTGGTATATGTCGCCGTG 1102
Db 1 CCAATGAGAGAGCGGAGCATATGAGTATGATGCCCCCTACTTGGTATATGTCGCCGTG 60
OY 1103 TCGCCCTTCGCTTGGAGAGCCCAAGATGGAAGAGTGTCTTTCAGATGGATGGGAATTGA 1162
Db 61 TCGCCCTTCGCTTGGAGAGCCCAAGATGGAAGAGTGTCTTTCAGATGGATGGGAATTGA 120
OY 1163 TGGTTAGGAGAGCGGCTGAGAGGCGAGTGCACCAACTACTTGTGATGTCAGCGGTT 1222
Db 121 TGGTTAGGAGAGCGGCTGAGAGGCGAGTGCACCAACTACTTGTGATGTCAGCGGTT 180
OY 1223 GCGTGGAGAGCGGCGCCGAG 1241
Db 181 GCGTGGAGAGCGGCGCCGAG 199

RESULT 11

US-10-040-862-2905
Sequence 2905, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther

FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-2905

Query Match 12.5% Score 197.4; DB 9; Length 199;
Best Local Similarity 99.5% Pred. No. 1.2e-45;
Matches 198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1043 CCATGAGAGAGCGGAGCATATGAGTATGATGCCCCCTACTTGGTATATGTCGCCGTG 1102
Db 1 CCAATGAGAGAGCGGAGCATATGAGTATGATGCCCCCTACTTGGTATATGTCGCCGTG 60
OY 1103 TCGCCCTTCGCTTGGAGAGCCCAAGATGGAAGAGTGTCTTTCAGATGGATGGGAATTGA 1162
Db 61 TCGCCCTTCGCTTGGAGAGCCCAAGATGGAAGAGTGTCTTTCAGATGGATGGGAATTGA 120
OY 1163 TGGTTAGGAGAGCGGCTGAGAGGCGAGTGCACCAACTACTTGTGATGTCAGCGGTT 1222
Db 121 TGGTTAGGAGAGCGGCTGAGAGGCGAGTGCACCAACTACTTGTGATGTCAGCGGTT 180
OY 1223 GCGTGGAGAGCGGCGCCGAG 1241
Db 181 GCGTGGAGAGCGGCGCCGAG 199

RESULT 12
US-10-015-219-658
Sequence 658, Application US/10015219
Publication No. US2003008299A1


```
NAME/KEY: misc_feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14: positive strandedness
NAME/KEY: misc_feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15: positive strandedness
NAME/KEY: misc_feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 16: positive strandedness
NAME/KEY: misc_feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17: positive strandedness
NAME/KEY: misc_feature
LOCATION: (70099)..(70662)
OTHER INFORMATION: ORF 18: positive strandedness
NAME/KEY: misc_feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19: positive strandedness
NAME/KEY: misc_feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20: negative strandedness
NAME/KEY: misc_feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21: negative strandedness
NAME/KEY: misc_feature
LOCATION: (75424)..(74213)
OTHER INFORMATION: ORF 22: negative strandedness
NAME/KEY: misc_feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23: positive strandedness
NAME/KEY: misc_feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24: negative strandedness
NAME/KEY: misc_feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25: negative strandedness
NAME/KEY: misc_feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26: negative strandedness
NAME/KEY: misc_feature
LOCATION: (81509)..(81682)
OTHER INFORMATION: ORF 27: negative strandedness
NAME/KEY: misc_feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28: negative strandedness
NAME/KEY: misc_feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29: positive strandedness
NAME/KEY: misc_feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30: positive strandedness
NAME/KEY: misc_feature
LOCATION: (85556)..(88845)
OTHER INFORMATION: ORF 31: positive strandedness
NAME/KEY: misc_feature
LOCATION: (85548)..(86803)
OTHER INFORMATION: ORF 32: positive strandedness
NAME/KEY: misc_feature
LOCATION: (87544)..(88420)
OTHER INFORMATION: ORF 33: positive strandedness; N-terminus only
US-09-937-039-1

Query Match 2.8% Score 44.8: DB 9: Length 88421:
Best Local Similarity 45.3% Pred. No. 0.02:
Matches 163: Conservative 0: Mismatches 197: Indels 0: Gaps 0:
```

```
215 AGCGCTTGCAGCTCTTCGCGAGTCAGCTGAGCCCTTTTGGCTGAGCTGAATGCCCT 274
16106 GCGAGCTACTGACGAGGAGTACCTTCATACGCGAGAACCGAGCCGCGCTCTCGCC 16047
275 TCAGCTGATGCTCAGTCAGCGCGGAGAACCGCCGAGAGCTGCTGGAGAGAC 334
16046 TCGCCCATACCTGACGAGCGAGCCGCGAGAGATCGAGCTCGACACTGCGCTGATC 15987
335 TGGGCGCTGGGAGCTCTGCTGCTCATGCTGAGAGAGCGCTGATGACAGAGTGTCA 394
15986 TCAGCAGACTCGCCGACCTGATACAGCGGCTTCCGCGGCGAGCTGCTGCGCTGCA 15927
395 ACGAGCTCATGACGCGCTTCAGTCGAGACCGCATTCAGAGAGCCCTGTGTAGCTCC 454
15926 ACGCAGAGCTGGCGGCGCTGAGCGCGCAGCTCAGAGCTGCGCAGCAGCTCAGCTCC 15867

RESULT 15
US-09-737-149-3
Sequence 3: Application US/09737149
Patent No. US20020077466A1
GENERAL INFORMATION:
APPLICANT: Spademe, Steven K
APPLICANT: Spademe, Kerry E
APPLICANT: Shinkels, Richard A
APPLICANT: Shinkels, Richard A
APPLICANT: Spirek, Kimberly A
TITLE OR INVENTION: Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-620 CIP
CURRENT APPLICATION NUMBER: US/09/737,149
PRIOR APPLICATION NUMBER: 60/170,564
PRIOR FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: 60/173,165
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: 60/173,362
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: 60/174,544
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 60/174,404
PRIOR FILING DATE: 2000-01-04
PRIOR APPLICATION NUMBER: 60/174,962
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: 60/223,929
PRIOR FILING DATE: 2000-08-09
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 2898
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2898)
US-09-737-149-3

Query Match 2.8% Score 43.8: DB 10: Length 2898:
Best Local Similarity 47.3% Pred. No. 0.019:
Matches 166: Conservative 0: Mismatches 182: Indels 3: Gaps 1:
```




GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 09:07:01 ; Search time 2408 seconds
(without alignments)
10579.527 Million cell updates/sec

Title: US-09-937-060A-19

Perfect score: 1573
Sequence: 1 gccccacagccgcctgcg.....gccctgaaaaaaaaaaaaaa 1573

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estlba:*
2: em_estlum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estcov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1547.8	98.4	2133	11	BC014439 Homo sapi
2	878	55.8	1054	14	BM808698 AGENCOURT
3	845.8	53.8	1137	14	BM810136 AGENCOURT
4	814.6	51.8	892	13	B1860351 AGENCOURT
5	787.8	50.1	888	14	BO647377 AGENCOURT
6	734.8	46.7	785	12	BG680521 AGENCOURT

Result No.	Score	Query Match	Length	ID	Description
7	732.8	46.6	772	9	AT1769914
8	716.8	45.6	1100	14	BM916526
9	643	40.5	676	10	AW963415
10	636.4	40.5	702	10	BE464487
11	630.6	40.1	875	13	BM006005
12	629	40.0	793	12	BG678689
13	625.4	39.8	662	13	BM386957
14	622	39.5	638	9	AI972156
15	620.2	39.4	809	12	BG280830
16	612.8	39.0	671	10	AV694791
17	606.8	38.6	619	10	AM662445
18	601.8	38.3	659	12	BE891653
19	600	38.1	940	12	BG026631
20	597	38.0	606	14	BQ251614
21	591	37.6	999	14	BO675531
22	582.2	37.0	603	10	BE275818
23	572	36.4	1005	13	BM559257
24	563.2	35.8	852	10	BE274434
25	559	35.5	1075	13	BM557357
26	545.8	34.7	2145	11	BC011432
27	532.6	33.9	842	13	B1648186
28	526.2	33.5	947	13	B1415545
29	510.2	32.4	587	9	AU147968
30	477.8	30.4	720	14	BQ109456
31	476.2	30.3	534	12	BE904632
32	469.2	29.8	856	13	B1757316
33	468	29.8	478	12	BE740866
34	455.6	29.0	1134	14	BM808666
35	449.4	28.6	697	13	B1255900
36	446.4	28.4	1232	13	BM556915
37	441.8	28.1	1115	13	BM550039
38	439	27.9	447	9	AI042283
39	426.2	27.1	914	13	B1823329
40	411.2	26.1	416	13	BM471470
41	407.8	25.9	732	12	BE728564
42	401.4	25.5	733	14	BM934157
43	400.2	25.4	662	12	BE893504
44	397.6	25.3	412	9	AA639414
45	397.2	25.3	618	10	AW916618

ALIGNMENTS

RESULT 1
LOCUS BC014439 2133 bp mRNA linear HTC 19-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4871343, mRNA.
ACCESSION BC014439
VERSION BC014439.1 GI:117939597
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2133)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCRD/PTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butlerfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1054)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubi Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1979 row: g column: 03
High quality sequence stop: 7
High quality sequence stop: 666.
Location/Qualifiers
1..1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5471258"
/clone.lib="NIH-MGC_41"
/tissue.type="melanotic melanoma, cell line"
/lab.host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 177 a 317 c 340 g 220 t
ORIGIN

Query Match 55.8%; Score 878; DB 14; Length 1054;
Best Local Similarity 94.0%; Pred. No. 1.6e-173;
Matches 981; Conservative 0; Mismatches 50; Indels 13; Gaps 6;

QY 272 CCTTCAGCTGATGCTCAGTACGAGCGGGAACACGCGGAGCTGGTCGGGCGAGG 331
DB 1 CCTTCAGCTGATGCTCAGTACGAGCGGGAACACGCGGAGCTGGTCGGGCGAGG 60
QY 332 AGCTGGGCGCTGGAGCGCTCTGTGTCATGCTTGAGAGCGGCGTGTACAGAGCTGG 391
DB 61 AGCTGGGCGCTGGAGCGCTCTGTGTCATGCTTGAGAGCGGCGTGTACAGAGCTGG 120
QY 392 TGAACGGGCTCATGAGAGCGGCTGACTGGAGACCGCATCCAGAACCCCTGTGTAGCC 451
DB 121 TGAACGGGCTCATGAGAGCGGCTGACTGGAGACCGCATCCAGAACCCCTGTGTAGCC 180
QY 452 TCCAGAGAGCTGTGGCAACGCGCTGGAGCTTCCTTGAACCAATTAATGCGGATAGAGC 511
DB 181 TCCAGAGAGCTGTGGCAACGCGCTGGAGCTTCCTTGAACCAATTAATGCGGATAGAGC 240
QY 512 AGCTCAGCAATGAGAGCTTCCTTGAACCAATGAGAGCTATTGCTGTGCGCGGCTGCTGT 571
DB 241 AGCTCAGCAATGAGAGCTTCCTTGAACCAATGAGAGCTATTGCTGTGCGCGGCTGCTGT 300
QY 572 CACCCATGAACCTGCTGTCTGTGACACGCGCTTGCGGCGCTTCTCTGTGCTCA 631
DB 301 CACCCATGAACCTGCTGTCTGTGACACGCGCTTGCGGCGCTTCTCTGTGCTCA 360
QY 632 GCGTGGCGCTGAGCTCATGCTGATGAGAGCTAGAGAGTGAAGTATGCGGCTGG 691
DB 361 GCGTGGCGCTGAGCTCATGCTGATGAGAGCTAGAGAGTGAAGTATGCGGCTGG 420
QY 692 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGCTGCTGGAGCGCTTGCACCTACCGCG 751
DB 1 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGCTGCTGGAGCGCTTGCACCTACCGCG 751

DB 421 GGGAGATGCGCTTCACTCTGTGGGACCTTCTGCTGCTGGAGCGCTTGCACCTACCGCG 480
QY 752 GCGGACTGGCGCTACCTCCCTGTAGAGAGTGGGTTCCAGACACCTGCTCCCGCGTTG 811
DB 481 GCGGACTGGCGCTACCTCCCTGTAGAGAGTGGGTTCCAGACACCTGCTCCCGCGTTG 540
QY 812 TGGTCCAGCAGGCGCGGTAGATGACACCTTGTGCGACTGAGAGAGCCAGTCCCTTC 871
DB 541 TGGTCCAGCAGGCGCGGTAGATGACACCTTGTGCGACTGAGAGAGCCAGTCCCTTC 600
QY 872 ACTGAGACGTGTGCGCGGAGAGACTTGTGCTAGCTGCGGACTGCTGACCTGCACC 931
DB 601 ACTGAGACGTGTGCGCGGAGAGACTTGTGCTAGCTGCGGACTGCTGACCTGCACC 660
QY 932 TGGGAGATGATGCTTGTGCGACCCATGCGCGCGGTGCGAGCTGGGTCATGATCTGT 991
DB 661 TGGGAGATGATGCTTGTGCGACCCATGCGCGCGGTGCGAGCTGGGTCATGATCTGT 720
QY 992 TCTACGTGC-GGGCGGAGAGTGTGCGGCGCATGCTGCTGCGCCCTTCCGCGCATGAG 1050
DB 721 TCTACGTGC-GGGCGGAGAGTGTGCGGCGCATGCTGCTGCGCCCTTCCGCGCATGAG 780
QY 1051 AAGGGCAGCATATGAGATGAATGAAATGCGCTACTTGTGATATGTCGCGGCTTC 1110
DB 781 AAGGGCAGCATATGAGATGAATGAAATGCGCTACTTGTGATATGTCGCGGCTTC 840
QY 1111 CGCTGGAGCGCCAAAGATGGGAAAGTGTGCTTGCACCTGGAATGGGAAATGATGTTA 1168
DB 841 CGCTGGAGCGCCAAAGATGGGAAAGTGTGCTTGCACCTGGAATGGGAAATGATGTTA 900
QY 1169 GC--GAGGCGCGTGCAGGCGCCAGGTGACCCCAACTACTTGTGATGTC---AGCGTT 1222
DB 901 GCGGAGCGCGTGCAGGCGCGGAGGCGAGTGCACCACTATTTTGGAGAGGCCCAAGGGGTG 960
QY 1223 GCGTGGAGCGCCCGCGCC--AGCTGGAAGCCCGCAGCAGATGCC--ACCGCAGAGAGCCC 1278
DB 961 GGGTGGAGCGCCCGCGCCCGCAGCTGGAAACCCCGCAGATGCGCCAGAAAGAAAGCC 1020
QY 1279 TTATGACCCCTGGGCGCGCTGTG 1302
DB 1021 TTTTATACCCCTGGGCGCGCTGTG 1044

RESULT 3
BM810136
LOCUS BM810136 1137 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT_6579661 NIH-MGC_41 Homo sapiens cDNA clone IMAGE:547089
5', mRNA sequence.
ACCESSION BM810136
VERSION BM810136.1 GI:19126959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1137)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubi Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1976 row: f column: 10
High quality sequence stop: 623.
Location/Qualifiers
1..1137
/organism="Homo sapiens"

Db	181	66TGTCTGTACACCATATGAACCTGCTGTCTGTGCACAGAGGCTTG66GGCTGTGCGCTCTTCT	240
QY	623	CTGTGTACAGCCTGTGCGCTTG66GGCTTTCATTTGCTGATATGTGCACCTTAGAGATGAAGATATC	682
Db	241	CTGTGTCTACAGCCTGTGCGCTTG66GGCTTTCATTTGCTGATATGTGCACCTTAGAGATGAAGATATC	300
QY	683	GGCGCTCTGGGGGAATGGCTTTCACCTTG66GACCTTCTGCTGTGGCACCCTTGCGCA	742
Db	301	GGCGCTCTGGGGGAATGGCTTTCACCTTG66GACCTTCTGCTGTGGCACCCTTGCGCA	360
QY	743	CCCTACCGGCGGACCTGGCCCTACCTTCCCTGTAGGAGAAGTGGTTCGAAGACCTTGCT	802
Db	361	CCCTACCGGCGGACCTGGCCCTACCTTCCCTGTAGGAGAAGTGGTTCGAAGACCTTGCT	420
QY	803	CCCCGTTGTGTGCACAGAGGCGCGGTAGATGACACCTTGTGCCACTGTGAGAGGCAG	862
Db	421	CCCCGTTGTGTGCACAGAGGCGCGGTAGATGACACCTTGTGCCACTGTGAGAGGCAG	480
QY	863	TGCGCTCTCACTGTGACACATGCTGCCCCACAGAGACTTGTGTCTAGTCTTGACACTGCTGC	922
Db	481	TGCGCTCTCACTGTGACACATGCTGCCCCACAGAGACTTGTGTCTAGTCTTGACACTGCTGC	540
QY	923	ACTGGACACTGGGACAGTATGATGTTTGTGTGACACCATGGGCGGCTGTGCACCTGGGCGCA	982
Db	541	ACTGGACACTGGGACAGTATGATGTTTGTGTGACACCATGGGCGGCTGTGCACCTGGGCGCA	599
QY	983	TGCATCTGTTCTAGTGTGGGCGGCGGAGTGTCTGTGTCATGCTGTGCGCTCTTCTCG	1042
Db	600	TGCATCTGTTCTAGTGTGGGCGGCGGAGTGTCTGTGTCATGCTGTGCGCTCTTCTCG	659
QY	1043	CCATGAGAGAGGACAGCATATGAGATATGAAATGCCCCCTACTTGGTATATGTGCCCCGTG	1102
Db	660	ACATGAGAGAGGACAGCATATGAGATATGAAATGCCCCCTACTTGGTATATGTGCCCCGTG	719
QY	1103	TGCGCTTCCGCTTGGAGGCCAAGAGTGTG-GAAAGGTGTGTTGTCAGTGCAGTGGGGAATG	1161
Db	720	TGCGCTTCCGCTTGGAGGCCAAGAGTGTG-GAAAGGTGTGTTGTCAGTGCAGTGGGGAATG	779
QY	1162	ATGATTACGAGAGGCGCTGTCAAGGCGCAGGTGACACCAACTACTTCTGTGATGTCAGCGGT	1221
Db	780	ATGATTACGAGAGGCGCTGTCAAGGCGCAGGTGACACCAACTACTTCTGTGATGTCAGCGGT	838
QY	1222	TGCGTGAAGCCCCCGCCAGCTGGAAGGCCCCAGCAGATGCCACCGCAGAAGAGC	1276
Db	839	TGCGTGAAGCCCCCGCCAGCTGGAAGGCCCCAGCAGATGCCACCGCAGAAGAGC	891
RESULT 5			
BO647377			
LOCUS	BO647377	888 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT 8414975 NIH_MGC_100 Homo sapiens cdna clone IMAGE:62720711		
ACCESSION	BO647377		
VERSION	BO647377.1	GI:21771549	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 888)		
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: CGAP (Stanford)		
	cDNA Library Preparation: Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: L10CM2449	row: n	column: 08

FEATURES									
source									
High quality sequence start: 6									
High quality sequence stop: 668.									
Location/Qualifiers									
1. .888									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="IMAGE:6272071"									
/clone_1lb="NIH_MGC_100"									
/tissue_type="hepatocellular carcinoma, cell line"									
/lab_host="DH10B (phage-resistant)"									
/note="Organ: Liver; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming, Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(g). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."									
BASE COUNT 137 a 274 c 284 g 193 t									
ORIGIN									
Query Match 50.1%; Score 787.8; DB 14; Length 888;									
Best Local Similarity 97.1%; Pred. No. 1.2e-154;									
Matches 834; Conservative 0; Mismatches 22; Indels 3; Gaps 3									
QY	537	CAACGACACCGCTATTGCTGTGCCGCCGCGCTGTCACACCATGAACCTTCTCTGTGCA	596						
Db	30	CCACGACGACGCTATTGCTGTGCCGCCGCGCTGTCACACCATGAACCTCTCTCTGTGCA	89						
QY	597	CACGGCTTGGGGGGCTGCGCCTCTTCTCTGTGCTCAGCCTGGCCTGGGGCTTCATTGCTGA	656						
Db	90	CACGGCTTGGGGGGCTGCGCCTCTTCTCTGTGCTCAGCCTGGCCTGGGGCTTCATTGCTGA	149						
QY	657	TGTGACCTTAGAGAGTGAAGATATCGCGCTGTGGGGGAGATGCCCTTACTCTGGGACAC	716						
Db	150	TGTGACCTTAGAGAGTGAAGATATCGCGCTGTGGGGGAGATGCCCTTACTCTGGGACAC	209						
QY	717	CTTCTGTGAGTCTGGGAGCCCTTGGACCTACACGCGGAGCAGTGGCTTACTCTCTTAGG	776						
Db	210	CTTCTGTGAGTCTGGGAGCCCTTGGACCTACACGCGGAGCAGTGGCTTACTCTCTTAGG	269						
QY	777	AAGAGTGGGTTTCAAGACACCTTGCCCTCCCGCTTGTGTCCAGCAGGGCCCGGTAGATGC	836						
Db	270	AAGAGTGGGTTTCAAGACACCTTGCCCTCCCGCTTGTGTGTCCAGCAGGGCCCGGTAGATGC	329						
QY	837	ACACCTTGTGCCACTGAGAGAGGACCAATGCCCTCTCACTGAGACATGTGTGCCGACAGAGA	896						
Db	330	ACACCTTGTGCCACTGAGAGAGGACCAATGCCCTCTCACTGAGACATGTGTGCCGACAGAGA	389						
QY	897	CTTTTGTGCTAGTCTGGGACCTGCTGCACCTGAGACCTGGGGGACATGAGATGTTGCTGCACC	956						
Db	390	CTTTTGTGCTAGTCTGGGACCTGCTGCACCTGAGACCTGGGGGACATGAGATGTTGCTGCACC	449						
QY	957	CATGGGCGCGCTGTGACAGCTGGCGGTATGCTTCTTCAAGTGGCGGGCGGAGATGTCTCG	1016						
Db	450	CATGGGCGCGCTGTGACAGCTGGCGGTATGCTTCTTCAAGTGGCGGGCGGAGATGTCTCG	509						
QY	1017	TGGCATGCTGCTGGCGCTTCTCTTGCGCATGTGAGAGAGGGCAGGACATATGAGATATGAAG	1076						
Db	510	TGGCATGCTGCTGGCGCTTCTCTTGCGCATGTGAGAGAGGGCAGGACATATGAGATATGAAG	569						
QY	1077	CCCCACTTTGGTATATGTCCCGTGCTCCCTTCCCTTGGAGAGCCCAAGATATGGGAAAGG	1136						
Db	570	CCCCACTTTGGTATATGTCCCGTGCTCCCTTCCCTTGGAGAGCCCAAGATATGGGAAAGG	629						
QY	1137	TGTGTTTGCAGTGAATGGGGAATTGATGTATGAGCGAGCGCTGTCAGGCGCAGTGTGACCC	1196						
Db	630	TGTGTTTGCAGTGAATGGGGAATTGATGTATGAGCGAGCGCGCTGTCAGGCGCAGTGTGACCC	689						
QY	1197	AAACTACTTTCGTGATGTCAGC-GCTGGCGTGGAGCCCGCCGCCAG-CGTGGAAGGCCCCAG	1254						
Db	690	AAACTACTTTCGTGATGTCAGC-GCTGGCGTGGAGCCCGCCGCCAG-CGTGGAAGGCCCCAG	749						

High quality sequence stop: 467.

FEATURES

source

Location/Qualifiers
1..772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:240431"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pF73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1323912-1325831, 1471368-1472903 and 1492104-1493253). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 170 a 244 c 213 g 143 t 2 others
ORIGIN

Query Match 46.6%; Score 732.8; DB 9; Length 772;
Best Local Similarity 96.9%; Pred. No. 3.7e-143;
Matches 746; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 784 GGTCCAAACACCTGCTCCCTCCCTGTTGTTGTCAGCAGAGGCCGGGTAGATGACACCTT 843
DB 772 GTTTCAAGACACNTGCTTCCCTCCCTGTTGTTGTCAGCAGAGGCCGGGTAGATGACACCTG 713
OY 844 GTGCCACTGAGAGACCACTGCTCCCTGTTGTTGTCAGCAGAGGCCGGGTAGATGACACCTT 903
DB 712 GTGCCACTGAGAGACCACTGCTCCCTGTTGTTGTCAGCAGAGGCCGGGTAGATGACACCTT 653
OY 904 CTAGTCCCTGAGACCTGCTGACACCTGAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 963
DB 652 CTAGTCCCTGAGACCTGCTGACACCTGAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 593
OY 964 CCTGTGACGCTGAGGCTGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1023
DB 592 CCTGTGACGCTGAGGCTGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 533
OY 1024 CTGTGCGGCTTCTCTGCGCATGAGAGGCGAGCATATGATGATGATGATGATGATGATGATG 1083
DB 532 CTGTGCGGCTTCTCTGCGCATGAGAGGCGAGCATATGATGATGATGATGATGATGATGATG 473
OY 1084 TTGCTATATGTCGCGGCTTCTCTGCGCATGAGAGGCGAGCATATGATGATGATGATGATG 1143
DB 472 TTGCTATATGTCGCGGCTTCTCTGCGCATGAGAGGCGAGCATATGATGATGATGATGATG 413
OY 1144 GCAGTGAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1203
DB 412 GCAGTGAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 353
OY 1204 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1263
DB 352 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 293
OY 1264 CGGCGAAGAGAGGCTTATGACCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1323
DB 292 CGGCGAAGAGAGGCTTATGACCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 233
OY 1324 CCTTCCCT 1383
DB 232 CCTTCCCT 173
OY 1384 CT 1443
DB 172 CT 113
OY 1444 GTCT 1503
DB 112 GTCT 53

OY 1504 TTCTGAGACCCGCCACGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1553
DB 52 TTTTGAGACCCGCCACGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3

RESULT 8
BM916526
LOCUS BM916526 1100 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
5', mRNA sequence.
ACCESSION BM916526
VERSION BM916526.1 GI:193662905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1100)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: DCTP/DMP

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
plate: LLCM2009 row: n column: 15
High quality sequence stop: 625.

FEATURES

source

Location/Qualifiers
1..1100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5482958"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGGAGG. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others
ORIGIN

Query Match 45.6%; Score 716.8; DB 14; Length 1100;
Best Local Similarity 89.8%; Pred. No. 7.6e-140;
Matches 848; Conservative 0; Mismatches 75; Indels 21; Gaps 7;

OY 272 CCTTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
DB 1 CCTTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
OY 332 AGCTGAGCGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
DB 61 AGCTGAGCGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
OY 392 TGAAGGCTCATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
DB 121 TGAAGGCTCATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
OY 452 TCCACAGAGCTTGGCAACGGCTGCGACCTCTTGAACCTTATGCTGCTATGAGC 511
DB 181 TCCACAGAGCTTGGCAACGGCTGCGACCTCTTGAACCTTATGCTGCTATGAGC 240
OY 512 AGGTCAACATGAAGACCTCTGACCAACTGACGATATTGCTGCTGCTGCTGCTGCTGCTGCTG 571

Db 241 AGGACACATGAGACCTCTGACCACTGACAGCTATTGCTGCGCGCGCTGCTGT 300
 572 CACCCATGAACGCTGCTGTCACACAGCGCTTCGCGGCTTCCTCTCTCTGTCGTC 631
 Db 301 CACCAATGAACGCTGCTGTCACACAGCGCTTCGCGGCTTCCTCTCTCTGTCGTC 360
 632 GCCGCGCTGCGGCTTCATGCTGATGAGACCTAGAGAGATGAGATGAGATGAGATG 691
 Db 361 GCCGCGCTGCGGCTTCATGCTGATGAGACCTAGAGAGATGAGATGAGATGAGATG 420
 692 GGGAGATGCGCTTCATGCTGAGGACCTTCCTGCTGCTGAGCGCGCGACCTAGCGCG 751
 Db 421 GGGAGATGCGCTTCATGCTGAGGACCTTCCTGCTGCTGAGCGCGCGACCTAGCGCG 480
 752 GCGGATGCGCTTCATGCTGAGGACCTTCCTGCTGCTGAGCGCGCGACCTAGCGCG 811
 Db 481 GCGGATGCGCTTCATGCTGAGGACCTTCCTGCTGCTGAGCGCGCGACCTAGCGCG 540
 812 TGGTCAGACGAGCGCGCTGATGACACCTGCTGCTGAGCGCGCGACCTAGCGCG 871
 Db 541 TGGTCAGACGAGCGCGCTGATGACACCTGCTGCTGAGCGCGCGACCTAGCGCG 600
 872 ACTGACAGATGCTGCGCGAGGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931
 Db 601 ACTGACAGATGCTGCGCGAGGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 932 TGGCG--AGTGAATGTTTGTGCTGACACCTAGCGCG--CGCTGCTGAGCTGCTGCTGCTG 988
 Db 661 TGGCGAGNTGAGATGTTTGTGCTGACCAATGAGGCGCTGCTGCTGCTGCTGCTGCTG 720
 989 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
 Db 721 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 1043 CCAT--GGAGAAGGCGAGCATATGAG--TATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1094
 Db 781 NCCTTGGAGAAAGGCGAGCATATGAG--TATGATGCTGCTGCTGCTGCTGCTGCTGCTG 840
 1095 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150
 Db 841 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 1151 ATGGGAAATGATGTTAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
 Db 901 ATGGGAAATGATGTTAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944

RESULT 9
 AM963415/c 676 bp mRNA linear EST 01-JUN-2000
 LOCUS AM963415
 DEFINITION EST373488 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM963415
 VERSION AM963415.1 GI:8153251
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 Hedge, P., Ol, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt,
 J.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
 Quackenbush, J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@ligr.org
 Plate: 187

Seq primer: Reverse.
 FEATURES Location/Qualifiers
 source 1..676
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="lib-MAGE resequences, MAGH"
 /note="Vector: pBluescriptSkm"
 BASE COUNT 147 a 220 c 188 g 121 t
 ORIGIN

Query Match 40.9%; Score 643; DB 10; Length 676;
 Best Local Similarity 98.4%; Pred. No. 2, 2e-124;
 Matches 660; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

882 GGTGCGCGACGAGACTTGTGCTAG--TCTGCGACACTGCTGCTGCTGCTGCTGCTGCTG 940
 Db 671 GGGCGCCCGACGAGACTTGTGCTAGTCCGGCCATGCGGCTGCTGCTGCTGCTGCTGCTG 612
 941 AGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
 Db 611 AGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
 1001 GCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1060
 Db 551 GCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
 1061 ATATGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
 Db 491 ATATGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
 1121 CCAAGATGGAAGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
 Db 431 CCAAGATGGAAGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
 1181 AGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240
 Db 371 AGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
 1241 GCTGGAAGCGCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1300
 Db 311 GCTGGAAGCGCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
 1301 TGGCTGAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1360
 Db 251 TGGCTGAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192
 1361 AGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1420
 Db 191 AGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132
 1421 TTGGGGGAGACGAGCGCCGATGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1480
 Db 131 TTGGGGGAGACGAGCGCCGATGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 72
 1481 CCTATGTAAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1540
 Db 71 CCTATGTAAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12
 1541 AGTACATTC 1551
 Db 11 AGTACATTC 1

RESULT 10
 BE464487/c 702 bp mRNA linear EST 27-JUL-2000
 LOCUS BE464487
 DEFINITION h217510.x1 NCI-GAP-GC6 Homo sapiens cDNA clone IMAGE:3208290 3'
 similar to TR:088886 088886 SPHINGOSINE KINASE.; mRNA sequence.
 ACCESSION BE464487
 VERSION BE464487.1 GI:9510262
 KEYWORDS EST
 SOURCE human.
 ORGANISM Homo sapiens

[illegible]

Query Match	40.1%	Score 630.6;	DB 13;	Length 875;
Best Local Similarity	88.9%;	Pred. No. 8.2e-122;		
Matches 778; Conservative	0;	Mismatches 74;	Indels 23;	Gaps 8;

Oy 522 TGAACACTCCAGACCAACCTGACCGCTATGCTGCGCGCGGCTGCTGTACACCATGAA 581
 |||

Db 1 TGAAGACCTCTGACCACTGCAACGCTATGCTGTCGCCCGCCGCTGCTGACCCATGAA 60
Oy 582 CCTGCTGTCTGTGACACAGGCTTCGGGGCTGGCCCTCTCTCTGTGCTACAGCTGACCTG 641
Db 61 CCTGCTGTCTGTGACACAGGCTTCGGGGCTGGCCCTCTCTCTGTGCTACAGCTGACCTG 120
Oy 642 GCGCTTATGCTGATGTGACCTAGAGAGTGAAGATGAGGCTGCTGCGGGAGATGCG 701
Db 121 GCGCTTATGCTGATGTGACCTAGAGAGTGAAGATGAGGCTGCTGCGGGAGATGCG 180
Oy 702 CTTCACCTGCGGACCTTCTCTGCTGCGGACGCTTCCACCTACCGGCGGCGACTGCG 761
Db 181 CTTCACCTGCGGACCTTCTCTGCTGCGGACGCTTCCACCTACCGGCGGCGACTGCG 240
Oy 762 CTACCTGCTGTGAGGAAGAGTGGGTTCACAGACAGCTGCCCGCGTGTGCTGACAGA 821
Db 241 CTACCTGCTGTGAGGAAGAGTGGGTTCACAGACAGCTGCCCGCGTGTGCTGACAGA 300
Oy 822 GGGCGCGGTGATGACACCTGCTGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
Db 301 GGGCGCGGTGATGACACCTGCTGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy 882 GGTGCGCGGACGAGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941
Db 361 GGTGCGCGGACGAGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Oy 942 GATGCTGCTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001
Db 421 GATGCTGCTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 1002 GGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
Db 481 GGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Oy 1062 TATGCA-GTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
Db 541 TATGAGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 1120 CCGAAGATGAGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
Db 601 CCGAAGATGAGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Oy 1176 CGTGCAGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
Db 661 GTGAGAGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Oy 1231 CCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1282
Db 721 ACAGAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Oy 1283 GACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
Db 781 GACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Oy 1340 AGGCGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1374
Db 841 AGGCGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 874

RESULT 12
LOCUS B6678689 793 bp mRNA linear EST 01-MAY-2001
DEFINITION 60262448BFL NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749573 5',
mRNA sequence.
ACCESSION B6678689
VERSION B6678689.1 GI:13910086
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM10602 row: P column: 22
High quality sequence stop: 669.
FEATURES
source location/qualifiers
1..793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4749573"
/issue_lib="NCI_CGAP_Skn4"
/issue_type="squamous cell carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NCI;
Site_2: Sall; Cloned unidirectionally; primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP library."
BASE COUNT 129 a 244 c 255 g 165 t
ORIGIN
Query Match 40.0%; Score 629; DB 12; Length 793;
Best Local Similarity 93.8%; Pred. No. 1..8e-121;
Matches 744; Conservative 0; Mismatches 35; Indels 14; Gaps 8;
Oy 270 CTCTTACCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329
Db 1 CTCTTACCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Oy 330 GGAAGTGGGCGCTGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389
Db 61 GGAAGTGGGCGCTGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Oy 390 GGTGAAGGCGCTATGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449
Db 121 GGTGAAGGCGCTATGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Oy 450 CTTCCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509
Db 181 CTTCCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Oy 510 GCAGGTACCAATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
Db 241 GCAGGTACCAATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Oy 570 GTCACCCATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
Db 301 GTCACCCATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Oy 570 GTCACCCATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
Db 301 GTCACCCATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Oy 630 CAGCGTGGCGCTGGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688
Db 361 CAGCGTGGCGCTGGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Oy 689 TGGGGAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db 421 TGGGGAGATGCGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Oy 749 GCGGCGAGCTGGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
Db 481 GCGGCGAGCTGGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Oy 807 CGTTGTGCTCAGCAGAGGCGCGTATGATGCAACACTTTCG-CCTGCC 864
Db 541 CGTTGTGCTCAGCAGAGGCGCGTATGATGCAACACTTTCG-CCTGCC 600
Oy 865 CCTCTCAGTGCAGAGTGGTGGCGGAGAGACT-TTGCTGCTAGTCTGCTGCTGCTGCTGCA 923


```

Db      601  CCATCTCACTGACAGTGTGCTCCGACGACGACTATCTGTACTGCTGACACTGTGCA 660
Oy      924  CTCGCACCTGGG--CAAGAGATGTTTGTGTGACACCATGGCCGCTGTG---CAGCTGGC 978
Db      661  CTCGCACCTGGGAGAGTCAAGTCACTGTACTGTGAGCCCATGTGGCCGATGTGGACCTGGC 720
Oy      979  GTCATGATCTGTTTCTAGTGTGCGGCGGAGTGTCTGCTG---CCATGCTGTGCGCCTC 1035
Db      721  GTCATGATCTGTGTGTACTGTGCGGAGGAGATGTATGTATGCTGCTGCTGACGCGCTTC 780
Oy      1036 TTCTGTGCGCATGC 1048
Db      781  TTCTGTGCGCATGC 793

RESULT 13
BM386957/c 662 bp mRNA linear EST 17-JAN-2002
LOCUS      BM386957
DEFINITION UI-R-CNI-cj1-c-13-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone
ACCESSION  BM386957
VERSION    BM386957.1 GI:18187010
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 662)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
COMMENT    discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
CONTACT   Soares, MB
PROGRAM   Program for Rat Gene Discovery and Mapping
INSTITUTION University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
TEL       319 335 8250
FAX       319 335 9565
EMAIL     mssoares@blue.weeg.uiowa.edu
NOTE      The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a homofide poly A
            tail. cDNA library Preparation: M.B. Soares lab Clone distribution:
            clones will be available through Research genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=Yes.

```

FEATURES source

```

1. 662
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cj1-c-13-0-UI"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies)"
/note="Vector: pT7SD-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-CNI
library is a subtracted library derived from the following
pool of seven normalized rat libraries: normalized rat
seminal vesicles, normalized rat penis, normalized rat
bladder, normalized rat cervix, normalized rat brown
adipose, normalized rat fundus, and normalized rat
salivary gland. It was constructed according to the
procedure described by Bonaldo, Lennon & Soares (Genome
Research Genome 6: 791-806, 1996). For construction of
the CNI library, plasmid DNA from the pool of normalized
libraries was electroporated into competent bacteria for
the production of single-stranded circular DNA. This was
then used as a tracer in a subtractive hybridization with
a driver (PCR amplified inserts from a plasmid DNA template

```

preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NM excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AMW through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKH-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BDA, R-CA1-BH2 through R-CA1-BJF, R-CA1-BJR, R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLV, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BFM through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVY through R-CW0-BMP, R-CM0-BXN through R-CM0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AOY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYL, R-DC0-BZB-C, R-DA0-BYJ through R-DA0-BJP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DD0-CAY through R-DD0-CBA, R-DD0-BZR through R-DD0-CAJ, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAI. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BLM-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BMI. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkx-a-06-0-UI, bkx-a-11-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, bla-a-12-0-UI, bla-f-02-0-UI, bla-a-11-0-UI, blic-e-95-0-UI, bld-f-08-0-UI, bld-f-02-0-UI, bld-h-04-0-UI, bly-a-05-0-UI, bly-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

7AG_SEO=None found"

RESULT 15
BG280830 809 bp mRNA linear EST 21-FEB-2001
LOCUS 602401209F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4543395 5',
DEFINITION mRNA sequence.
ACCESSION BG280830 GI:13029767
VERSION BG280830.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The i.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the i.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LHCMI225 row: b column: 04
High quality sequence stop: 688.
location/Qualifiers
1. 809
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4543395"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(C). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 165 a 240 c 250 g 153 t 1 others
ORIGIN
Query Match 39.4%; Score 620.2; DB 12; Length 809;
Best Local Similarity 96.9%; Pred. No. 1.2e-119;
Matches 717; Conservative 0; Mismatches 14; Indels 9; Caps 8;
QY 823 GGGCCGATGATGACACCTGTCGACATGGAGGAGCCCTGCTCACTGGACAGTG 882
DB 7 GGGCCGATGATGACACCTGTCGACATGGAGGAGCCCTGCTCACTGGACAGTG 66
QY 883 GTGCCCGACGAGGACTTGTAGTCTGACCTGGACCTGGACCTGGACCTGGAG 942
DB 67 GTGCCCGACGAGGACTTGTAGTCTGACCTGGACCTGGACCTGGACCTGGAG 126
QY 943 ATGTTGCTGACCCATGGGCGCTGTGACAGCTGGCTCATGATCTGTTTACGTGG 1002
DB 127 ATGTTGCTGACCCATGGGCGCTGTGACAGCTGGCTCATGATCTGTTTACGTGG 186
QY 1003 GCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062
DB 187 GCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
QY 1063 ATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1122
DB 247 ATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 306
QY 1123 AAGGATGGAAAGGTGTTTGGAGTGGATGGAGATGATGATGATGATGATGATG 1182
DB 307 AAGGATGGAAAGGTGTTTGGAGTGGATGGAGATGATGATGATGATGATGATG 366

QY 1183 GCGCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 1242
DB 367 GCGCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 426
QY 1243 TGAAGCCCCAGCAGATGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302
DB 427 TGAAGCCCCAGCAGATGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 1303 CCTAGTGTCTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1362
DB 487 CCTAGTGTCTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 1363 CTCTGTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1418
DB 547 CTCTGTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 1419 CTTCG-GGGGACAGGCCA-GAATGAAGTCTGAGTCAAGAGAGAGAGAGAGAG 1476
DB 607 CTTCGCGGGGACAGGCCAGCAATGAAGTCTGAGTCAAGAGAGAGAGAGAGAG 666
QY 1477 -GCTGCTATGTAGGCTTCTAG-TTGTCTGAGAGC-CCACCCGACGACCAATC 1533
DB 667 GCGTGCCTATGTAGGCTTCTAG-TTGTCTGAGAGC-CCACCCGACGACCAATC 726
QY 1534 CAATTAAGTGCATTCOCA 1553
DB 727 CAATTAAGTGCATTCOCA 746

Search completed: May 17, 2003, 14:17:36
Job time : 2441 secs



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:52:54 ; Search time 90 Seconds
(without alignments)
879.134 Million cell updates/sec

Title: US-09-937-060a-5
Perfect score: 2016
Sequence: 1 MDPAGGPRGVLPKRCRVLLV.....CVEPPSWKPOQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2009	99.7	384	4	Q96HV8
2	1999	99.2	398	4	Q96GK1
3	1644	81.5	329	6	Q9N0A5
4	1630.5	80.9	381	11	Q912N3
5	1630.5	80.9	388	11	Q912N3
6	1627.5	80.7	504	11	Q912N3
7	1599.5	79.3	383	11	Q912N3
8	1508	74.8	290	4	Q9B7G7
9	886.5	44.0	634	4	Q9H0Q2
10	886.5	44.0	634	4	Q9H0Q2
11	608	30.2	384	11	Q9D4D1
12	562	27.9	641	5	Q9VYH8
13	521.5	25.9	907	5	Q9VZM0
14	521	25.8	661	5	Q9VZM0
15	451	22.4	1240	10	Q9VZM0
16	411	20.4	624	5	Q9VZM0

17	374.5	18.6	473	5	Q18425
18	363.5	18.0	354	4	Q9NMU7
19	356.5	17.7	687	3	Q06147
20	355	17.6	458	3	Q14159
21	345	17.1	624	3	Q12246
22	272	13.5	763	10	Q9LRB0
23	265	13.1	757	10	Q94HY9
24	264.5	13.1	481	4	Q9BYH3
25	264.5	13.1	537	4	Q9BYH3
26	250.5	12.4	549	5	Q9TZ11
27	250	12.4	732	10	Q9FHL3
28	239	11.9	487	5	Q95S15
29	239	11.9	596	5	Q9VNA6
30	227.5	11.3	326	4	Q9UG55
31	213.5	10.6	422	4	Q9NPA8
32	203	10.1	421	11	Q9ESW4
33	185	9.2	533	10	Q9LU45
34	162.5	8.1	406	5	Q9VLJ0
35	155.5	7.7	306	16	Q927Y6
36	148	7.3	173	11	Q9D087
37	140.5	7.0	306	16	Q9Y497
38	140.5	7.0	338	10	Q94C71
39	136.5	6.8	309	16	Q34799
40	131	6.5	586	10	Q949C3
41	130	6.4	295	16	Q9KBM4
42	129.5	6.4	303	16	Q31502
43	127	6.3	342	16	Q9CTE3
44	126	6.2	302	16	Q9CEN8
45	125	6.2	340	16	Q9A0J4

ALIGNMENTS

RESULT 1	ID	Q96HV8	PRELIMINARY:	PRT:	384 AA.
AC	Q96HV8:				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Unknown (protein for MGC:15041).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=SKIN;				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: BC008040; AAH08040.1;				
DR	InterPro: IPR001395; Aldo/ket_red.				
DR	InterPro: IPR001206; DAGC.				
DR	InterPro: IPR003622; DAG_kin_cat.				
DR	Pfam: PF00781; DAGC: 1.				
DR	ProDom: PD005043; DAG_kin_cat; 1.				
DR	PROSITE: PS00063; ALDOKEETO_REDUCTASE_3; UNKNOWN_1.				
SO	SEQUENCE 384 AA; 42474 MW; P82999FP306113B0 CRC64;				
Query Match	99.7%;	Score 2009;	DB 4;	Length 384;	
Best Local Similarity	99.7%;	Pred. No. 5e-168;			
Matches	383;	Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;
Qy	1	MDPAGGPRGVLPKRCRVLLVLPNPGKGAQLQLEFSHVQPLLAELISFTLMLTERRNA	60		
Db	1	MDPAGGPRGVLPKRCRVLLVLPNPGKGAQLQLEFSHVQPLLAELISFTLMLTERRNA	60		
Qy	61	RELVAESELGNDALVAVSGGLMHEVYVNGLMERPDMEATQKPLCSLPASSGNALASL	120		
Db	61	RELVAESELGNDALVAVSGGLMHEVYVNGLMERPDMEATQKPLCSLPASSGNALASL	120		
Qy	121	NHYAGYEQVTNEDLLTNCILLICRRLSLSPMNLISHTASGLRLFVLSIAMGFADVIDLE	180		

```

Db 121 NHVAGDYNTNEDLLTNCITLLCRLLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
OY 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
Db 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVHLLFYVRAGVSRAML 300
Db 241 LEEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVHLLFYVRAGVSRAML 300
OY 301 RLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVSBAVVOGVHNPYFM 360
Db 301 RLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVSBAVVOGVHNPYFM 360
OY 361 MVSGCEPPPSMKPQOMPPEEP 384
Db 361 MVSGCEPPPSMKPQOMPPEEP 384

```

RESULT 2

```

O96GK1 PRELIMINARY: PRT: 398 AA.
ID O96GK1
AC O96GK1:
DT 01-DEC-2001 (Tremblrel, 19, Created)
DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel, 20, Last annotation update)
DE Unknown (protein for MGC:15040).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC009419; AAI09419.1;
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00063; ALDOKEO_REDUCTASE_3; UNKNOWN_1.
DR PROSITE: PS00063; ALDOKEO_REDUCTASE_3; UNKNOWN_1.
SO SEQUENCE 398 AA; 43944 MW; 7F2C9B26C030E560 CRC64;

```

Query Match 99.2%; Score 1999; DB 4; Length 398;

Best Local Similarity 96.5%; Prid. No. 3.9e-107; Matches 384; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

```

OY 1 MDP-----ACGRCVLPKPCRYVLYLNPGRGKALQIFRSHVOPLLAEAE 46
Db 1 MDPVVGCGRGLEGFVPSGGRGVLPKPCRYVLYLNPGRGKALQIFRSHVOPLLAEAE 60
OY 47 ISTTLLTERRNAHRELVRSEELGRMDALVYVSGDGLMHEVYVNGLMRPMETALOKPJC 106
Db 61 ISTTLLTERRNAHRELVRSEELGRMDALVYVSGDGLMHEVYVNGLMRPMETALOKPJC 120
OY 107 SLPASGNAALASLNHYAGYEOVTNEDLLTNCITLLCRLLSPNNLSLHTASGLRLEFSV 166
Db 121 SLPASGNAALASLNHYAGYEOVTNEDLLTNCITLLCRLLSPNNLSLHTASGLRLEFSV 180
OY 167 LSLANGFIADVDLESEKRRLEGRMFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASP 226
Db 181 LSLANGFIADVDLESEKRRLEGRMFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASP 240
OY 227 VVVOGQPVDAHLVLEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVH 286
Db 241 VVVOGQPVDAHLVLEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVH 300
OY 287 LTVYRAGVSRAMLRLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVS 346
Db 301 LTVYRAGVSRAMLRLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVS 360

```

```

OY 347 SEAVOGVHNPYFMVSGCEPPPSMKPQOMPPEEP 384
Db 361 SEAVOGVHNPYFMVSGCEPPPSMKPQOMPPEEP 398

```

RESULT 3

```

O9N0A5 PRELIMINARY: PRT: 329 AA.
ID O9N0A5
AC O9N0A5:
DT 01-OCT-2000 (Tremblrel, 15, Created)
DT 01-OCT-2000 (Tremblrel, 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel, 20, Last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN SEQUENCE FROM N.A.
RP Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB046025; BAB01607.1;
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKc; 1.
DR PROSITE: PS00063; ALDOKEO_REDUCTASE_3; UNKNOWN_1.
DR PROSITE: PS00063; ALDOKEO_REDUCTASE_3; UNKNOWN_1.
SO SEQUENCE 329 AA; 36393 MW; 7F430AB0CA5FDC7 CRC64;

```

Query Match 81.5%; Score 1644; DB 6; Length 329;

Best Local Similarity 98.1%; Prid. No. 4.4e-136; Matches 316; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 MDPAGGRCVLPKPCRYVLYLNPGRGKALQIFRSHVOPLLAEAEISTTLLTERRNAH 60
Db 1 MDPAGGRCVLPKPCRYVLYLNPGRGKALQIFRSHVOPLLAEAEISTTLLTERRNAH 60
OY 61 RELVRSBELGRMDALVYVSGDGLMHEVYVNGLMRPMETALOKPJCISLPASGNAALASL 120
Db 61 RELVRSBELGRMDALVYVSGDGLMHEVYVNGLMRPMETALOKPJCISLPASGNAALASL 120
OY 121 NHVAGDYNTNEDLLTNCITLLCRLLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
Db 121 NHVAGDYNTNEDLLTNCITLLCRLLSPNNLSLHTASGLRLEFSVLSLAWGFIADVDLE 180
OY 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
Db 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVSGSKTPASPVVVOGQPVDAHLVP 240
OY 241 LEEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVHLLFYVRAGVSRAML 300
Db 241 LEEPPVSHWTVPDEDFVLYLALSHLSGSEMFAPMGRCAAGVHLLFYVRAGVSRAML 300
OY 301 RLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVSBAVVOGVHNPYFM 360
Db 301 RLFLAMEGRHMEYECPLYVAVFAFLRPEKDGKGFVAVDGLMVSBAVVOGVHNPYFM 360

```

RESULT 4

```

O91ZN3 PRELIMINARY: PRT: 381 AA.
ID O91ZN3
AC O91ZN3:
DT 01-DEC-2001 (Tremblrel, 19, Created)
DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Sphingosine kinase 1a.

```

[illegible][illegible]

Best Local Similarity 81.4%; Pred. No. 2,1e-134;
Matches 307; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

QY 7 PRGVLPKPCRYLVLLNPGCGKALQIFRSVOPPLAEAEISFTLTERRNHARELVAS 66
DB 129 PRGLPRLPCRYLVLLNPGCGKALQIFRSVOPPLAEAEISFTLTERRNHARELVAS 188
QY 67 EELGRMDALVMSGDLMEHVEVNGLMERDMEATOKPLCSLPGSGNALAASVINYAGY 126
DB 189 EELGRMDALVMSGDLMEHVEVNGLMERDMEATOKPLCSLPGSGNALAASVINYAGY 248
QY 127 EGYTNEEDLITNCTLLCRLLSPMNLISHTASGLRFSVSLMGEFLAVDSEKRYR 186
DB 249 EGYTNEEDLITNCTLLCRLLSPMNLISHTASGLRFSVSLMGEFLAVDSEKRYR 308
QY 187 LGEEMFTLGTFLRLAALRTYRGRLATLPYRGVSGKTPASPVYVQGGPYDAHLVPLEPVP 246
DB 309 LGEIRFTVGTFFRLASLRTYRGRLATLPYRGVSGKTPASPVYVQGGPYDAHLVPLEPVP 367
QY 247 SHWTVVPEDEFLVYLLALSHLSGSEMFAPMGRCAGVYMHLEFVYRAGVSRAMLRLFLAM 306
DB 368 SHWTVVPEDEFLVYLLALSHLSGSEMFAPMGRCAGVYMHLEFVYRAGVSRAMLRLFLAM 427
QY 307 EKGRIHMEYECPLYVYVVAFFLEPKDGKGVFAVDGELMVSFAVQGVHNPYFMVYSGCV 366
DB 428 OKGRIHMEYECPLYVYVVAFFLEPKDGKGVFAVDGELMVSFAVQGVHNPYFMVYSGCV 487
QY 367 EPPPSWKPOQMPPEEP 383
DB 488 DAPSGRSRGPPEEP 504

RESULT 7

Q91V26 PRELIMINARY: PRT: 383 AA.
AC 091V26: 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sphingosine kinase 1f (Sphingosine kinase 1a) (Sphingosine kinase 1c)
GN SPHK1F OR SPHK1C OR SPHK1D OR SPHK1E.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11)
RP SEQUENCE FROM N.A.
RA Imanura T., Ohgawa J., Ito S., Ogawa T., Hattori N., Tanaka S.,
RA Shioita K.;
RT *Cpg island of rat sphingosine kinase-1 gene: tissue-dependent DNA
RT methylation status and multiple alternative first exons.";
RL Genomics 78:117-125(2001).
DR EMBL: AB049575; BAB62324.1; -
DR EMBL: AB049575; BAB62320.1; -
DR EMBL: AB049575; BAB62321.1; -
DR EMBL: AB049575; BAB62322.1; -
DR EMBL: AB049575; BAB62323.1; -
DR EMBL: AB049575; BAB62324.1; -
DR InterPro: IPR003622; DAG_kin_cat.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGC; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
KW kinase.
SQ SEQUENCE 383 AA; 42418 MW; CAAD5817BF4B9507 CRC64;

Query Match 79.3%; Score 1599.5; DB 11; Length 383;
Best Local Similarity 78.1%; Pred. No. 4.3e-132;
Matches 300; Conservative 42; Mismatches 41; Indels 1; Gaps 1;
QY 1 MDPAGPGRVLPKPCRYLVLLNPGCGKALQIFRSVOPPLAEAEISFTLTERRNH 60
DB 1 MDPAGPGRVLPKPCRYLVLLNPGCGKALQIFRSVOPPLAEAEISFTLTERRNH 60

QY 61 RELVASEELGRMDALVMSGDLMEHVEVNGLMERDMEATOKPLCSLPGSGNALAASL 120
DB 61 RELVASEELGRMDALVMSGDLMEHVEVNGLMERDMEATOKPLCSLPGSGNALAASL 120
QY 121 NHVAGYDGVNEDLITNCTLLCRLLSPMNLISHTASGLRFSVSLMGEFLAVDSEKRYR 180
DB 121 NHVAGYDGVNEDLITNCTLLCRLLSPMNLISHTASGLRFSVSLMGEFLAVDSEKRYR 180
QY 181 SEXYRRLGEMFTLGTFLRLAALRTYRGRLATLPYRGVSGKTPASPVYVQGGPYDAHLV 240
DB 181 SEXYRRLGEMFTLGTFLRLAALRTYRGRLATLPYRGVSGKTPASPVYVQGGPYDAHLV 240
QY 241 LEEYPSLWTVVPEDEFLVYLLALSHLSGSEMFAPMGRCAGVYMHLEFVYRAGVSRAML 300
DB 241 LEEYPSLWTVVPEDEFLVYLLALSHLSGSEMFAPMGRCAGVYMHLEFVYRAGVSRAML 300
QY 301 RLFLAMEKGRHMEYECPLYVYVVAFFLEPKDGKGVFAVDGELMVSFAVQGVHNPYFM 360
DB 301 RLFLAMEKGRHMEYECPLYVYVVAFFLEPKDGKGVFAVDGELMVSFAVQGVHNPYFM 360
QY 361 MVSGCDEPPPSWKPOQMPPEEP 384
DB 361 MVSGCDEPPPSWKPOQMPPEEP 384

RESULT 8

Q9BTG7 PRELIMINARY: PRT: 290 AA.
AC 09BTG7: 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to sphingosine kinase 1 (fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC Tissue-Skin;
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC004112; AA04112.1; -
DR EMBL: BC004112; AA04112.1; -
KW kinase.
FT NON_TER
SQ SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;

Query Match 74.8%; Score 1508; DB 4; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 EFAIOLKPLCSLPGSGNALAASVINYAGYEDVTNEDLITNCTLLCRLLSPMNLISHT 157
DB 4 EFAIOLKPLCSLPGSGNALAASVINYAGYEDVTNEDLITNCTLLCRLLSPMNLISHT 157
QY 158 ASGLRLFSVSLMGEFLAVDSEKRYRLEGEMFTLGTFLRLAALRTYRGRLATLPYGR 217
DB 64 ASGLRLFSVSLMGEFLAVDSEKRYRLEGEMFTLGTFLRLAALRTYRGRLATLPYGR 217
QY 218 VGSKTPASPVYVQGGPYDAHLVPLEPVSHTWTVVPEDEFLVYLLALSHLSGSEMFAPM 277
DB 218 VGSKTPASPVYVQGGPYDAHLVPLEPVSHTWTVVPEDEFLVYLLALSHLSGSEMFAPM 277
QY 278 GRCAGVYMHLEFVYRAGVSRAMLRLFLAMEKGRHMEYECPLYVYVVAFFLEPKDGK 337
DB 278 GRCAGVYMHLEFVYRAGVSRAMLRLFLAMEKGRHMEYECPLYVYVVAFFLEPKDGK 337
QY 338 FAVDGEMLVSEAVQGVHNPYFMVYSGCVPEPPPSWKPOQMPPEEP 384
DB 338 FAVDGEMLVSEAVQGVHNPYFMVYSGCVPEPPPSWKPOQMPPEEP 384

RESULT 9

[illegible]

RESULT	10
Q9BRN1	
ID	Q9BRN1
AC	Q9BRN1;
DT	01-JUN-2001 (TREMBlrel. 17, Created)
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE	Hypothetical 69.2 kDa protein.
OS	Homo sapiens (human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxId=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lymph;
RA	Strausberg R.;
RL	Submitted (APR-2001) to the EMBL/Genbank/DBDJ databases.
DR	EMBL: BC006161; AAH06161.1; -
DR	InterPro: IPR001206; DAgC.
DR	InterPro: IPR003622; DAg_Kin_cat.
DR	Pfam: PF00781; DAgC_1.
DR	ProDom: PD05043; DAg_Kin_cat; 1.
DR	SMART: SM00046; DAgC; 1.
KW	Hypothetical protein.
SO	SEQUENCE 654 AA; 69217 MW; F73FFCEC9300DA50F CRC64;

[illegible]

01-JUN-2001 (TREMBLER, 17, last sequence update)
 01-MAR-2002 (TREMBLER, 20, last annotation update)
 DE Adult male testis cDNA, RIKEN full-length enriched library.
 DE Clone:4933402E20, full insert sequence.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RX MEDLINE=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Ishii Y.,
 RA Atakawa T., Hara A., Fukuishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glassl C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bolunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Segal T., Shibata T., Storch K.-F.,
 RA Suzuki H., Togo O.K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Watanabe-Boliz A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA Functional annotation of a full-length mouse cDNA collection.
 RL Mature 403:685-690(2001).
 DR EMBL: AK016616; BAB0339.1;
 SO SEQUENCE 384 AA; 39865 MW; 6C94D1D3806092A3 CRC64;
 Query Match 30.2%; Score 608; DB 11; Length 384;
 Best Local Similarity 36.4%; Pred. No. 4, 3e-45;
 Matches 141; Conservative 38; Mismatches 88; Indels 120; Gaps 5;
 QY 104 PLCSLPGAGSALAAASINHYAGYQVNTNEDLTNCTLLCRRLSPMNLSTHTASGLRL 163
 DB 2 PLGVLPGSSGALAAAGVNHGFEQVGVLDLNCSTLCSGSPRLDLSVTLASGRC 61
 QY 164 FSVLSLAWGFIDVDESEKRYRGLSEKRPFLGTPFLALRTGRRGLATLVV----- 215
 DB 62 FSVLSLAWGFIDVDESEKRYRGLSEKRPFLGTPFLALRTGRRGLATLVV----- 215
 QY 216 ----- 215
 DB 122 PCHSLPRAKSELVLAAPARAPATHSPHRSVSDLPRLPQALVSPGSPRLPDLSLNG 181
 QY 216 -----GRGS-----KTFAS-----PVV 228
 DB 182 GRELTGDMGAGDAPLSPDLSPSSPMALKTAQSLFAGRPPEMAGSGRLPRTSHAPR 241
 QY 229 VQGGGVDAHLVPLEBVRPSKTVVPRDEPVVLALSHLSGSEMFARPMSCAAGVNL 288
 DB 242 STWGVDAHLVPLEBVRPSKTVVPRDEPVVLALSHLSGSEMFARPMSCAAGVNL 300
 QY 289 YVAGVSRAMLLRLFLMEKGRMEYECRPLVYVVAFRLEPKDKGVAFVDELVSE 348
 DB 301 VWRSGISRALRLITLAMEGNHNSLSCPHLGYAAAFARLEPRLPRLDVLDELVEG 360
 QY 349 AVQGVNRYFMVNSGCVPEPSKRP 375
 DB 361 PIAQVNHGRLATLTG---PAGCKRP 383

01-MAY-2000 (TREMBLER, 13, last sequence update)
 01-OCT-2001 (TREMBLER, 18, last annotation update)
 DE CG1747 protein.
 GN CG1747.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandus P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berns P., Bernier J., Brinkman P., Brottier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Gentler A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de la Cruz A., Deng Z., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hartis N.L., Harvey D., Heiman T.J., Hernandez J., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
 RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervin G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spiet E., Spradling A.C., Stapleton M., Strouf R., Sun E.,
 RA Switskas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
 RA Ye J., Yen R.F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong E.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Zuo R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RL The genome sequence of Drosophila melanogaster.
 RL Science 287:2185-2193(2000).
 DR EMBL: AB003486; AAF48045.1;
 DR FLYbase: FBgn0030300; CG1747.
 DR InterPro: IPR001206; DAGC.
 DR InterPro: IPR003622; DAGC_kin_cat.
 DR Pfam: PF00781; DAGC_kin_cat.
 DR ProDom: PD005043; DAGC_kin_cat; 1.
 DR SMART: SM00046; DAGC_kin_cat; 1.
 SO SEQUENCE 641 AA; 71421 MW; 1CE24E48DF14758D CRC64;
 Query Match 27.9%; Score 562; DB 5; Length 641;
 Best Local Similarity 32.5%; Pred. No. 9, 1e-41;
 Matches 148; Conservative 67; Mismatches 121; Indels 120; Gaps 16;
 QY 12 PRPC--RVLVLLNPRGKGKALQFPRSHVQPLLAFAEISPLMLTERNRNARLVSEF- 68
 DB 184 PADCGKQLLLNPKSSGKRELFQKVAAPLITAEVQYDLOITTPVAKFVTRRD 243
 QY 69 -LGRDVALVNSGGLMHEVVGMLERPRMPTAIOK-PLCSLPGAGSALAAASLNTAG- 125
 DB 244 LTRYSGLTVASGGLFYEVLNGLMERMQWRACREPLGLITPGSSNGSLAKSVANHCNE 303
 QY 126 -YEQVNTDLNCTLLCRRLSPMNLSTHTASGLR--LFSVLSLAWGFLADVDES 181

```

Db 304 PYEKRP-----IIHATLTCMAKSGSTPMDDVVRRLATROKHFWYSLVSGVGLIADIDIES 359
OY 182 EKYRILGEMRFTLCTFLTAIRYGRGLAVPYGRVSGKTPASPV----- 227
Db 360 ERLSIGMORTLWAIKRLIGLRKRGVSTL-LGK-GKRP--PVEARELPASTAAG 415
OY 228 -----VVQO----- 231
Db 416 IRSSLPLNAGEFHDLPBEDEGEAVLDGEQADAIISLRSVYRQHADSWMHMSMRTAYVS 475
OY 232 --GP-----VDA-----HLVPLEEVPVSR--WTVPVDEDFVLYL 261
Db 476 IGGPSMRSNRSMSISORIEAAMAEFAERVPPTGTLPLQMPLLSSDGM-ICEDGGFVAVH 534
OY 262 ALLHSIGSEMEAPMGRCAGVAMHLEFVRAGVSRAMLRLFLAMEKGRHMEY-ECPTLYV 320
Db 535 AAYTHHSSDVFAPESLDDGLLYLTIIRKGVSRHOLLNMLNMASTHLP.IGEDPTIK 594
OY 321 VVPVAFRIEPRKDGKGFVAVDGLVSEAVOGVHP 356
Db 595 VVPCRAFRIEPPSSDGLVVDGEREYGPICAFVMP 630

RESULT 13
OYVZWO PRELIMINARY: PRT: 907 AA.
AC O9VZWO
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE CG2159 protein.
GN Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OY NCBI_TaxID=7227; [1]
RN RN
RP SEQUENCE FROM N.A.
RX STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Aamundnes P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.H., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jajali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazkolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

```

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003476; AAF47706.1; -.
DR FlyBase; FBgn0035391; CG2159.
DR InterPro; IPR001251; CRAL_TRIO.
DR InterPro; IPR001206; DAGKC.
DR InterPro; IPR003622; DAG_Kin_cat.
DR Pfam; PF00650; CRAL_TRIO.1.
DR Pfam; PF00781; DAGKC.1.
DR ProDom; PD005043; DAG_Kin_cat.1.
DR SMART; SM00046; DAGKC.1.
DR SMART; SM00516; SEC14.1.
SO SEQUENCE 907 AA; 102561 MW; 58206C8E12A3F195 CRC64;

Query Match 25.9%; Score 521.5; DB 5; Length 907;
Best Local Similarity 30.5%; Pred. No. 5,1e-37;
Matches 134; Conservative 61; Mismatches 138; Indels 107; Gaps 8;

OY 16 RVLVLNPRGKGKALQLFRSHVQPLAEAEISFTLMLTERNNHARELYRSEELGRDAL 75
Db 216 RVLVLNPKSGSGDAREVFNHVTPLNEAEVPLDYVTKHSNFAIEFLSTRCLDAMCCV 275
OY 76 VVMSDGLMHEVNVNLMRPWETAI-QKPLSLPAGSGNLAASLNINAYAGVQYTNEDL 134
Db 276 VAVGDDGFEHEIVNGLORODAHVLPILALGIIPCGSGNGLARSIH-----CYNRPV 329
OY 135 LTNCITLLCRLLSPMNLISLHTASGLRFLSVLSLAWGFIADVDLESERKRLGEMRFTL 194
Db 330 LGALTYISGR--SSPDVYRVQDQS-RSLYSPLSLIGKGLISVNDLESERKRLGQRTFV 387
OY 195 GTFLALALRTYGRGLAVLPVGRVSGKTPASPVVVOGCV-----DAHLVPLEEVPVSH 248
Db 388 WLYRLVNLRTYNGRISYLLTFDHEVSTHSATGYAORRMOSRSCNTHIDMLNCPAPLY 447
OY 249 -----WT----- 250
Db 448 HSSAEYLPQEFADVISLETSLNOSFRSRCDVLSGSGSRSPYSISESLYHSLADESEFA 507
OY 251 -----VYDEDFVVLALLSHLSCSEMAFPM 277
Db 508 GLAASLNRQONTGPASFLPPLNLEPSEDCWLVBEGEFVMAHAYQTHLIGDHPAPK 567
OY 278 GRCAGVNHLEFVRAGVSRAMLRLFLAMEKGRHM-EYECPTLYVVPVAFRIEPRKDGK 336
Db 568 AOLNCGTYILLIRAGISPHLLSFLYNNMSSCTHLPESHDDHVKLVPAFRIEPRYDNHG 627
OY 337 VFAVVGELMAVSEAVOGVHP 356
Db 628 IITVDGERVEFGPLDAEVLPL 647

RESULT 14
OYVZWO PRELIMINARY: PRT: 661 AA.
AC O8T0C1
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD11247P.
GN CG2159.
OC Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OY NCBI_TaxID=7227; [1]
RN RN
RP SEQUENCE FROM N.A.
RX STRAIN-BERKELEY;

```

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guatin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nino J., Pacheco J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY069417; AAL3562.1;
DR SDB: 661 AA; 73870 MW; 4266219IEFB2D08 CRC64;
SQ

Query Match 25.88; Score 521; DB 5; Length 661;
Best Local Similarity 30.28; Pred. No. 3.8e-37;
Matches 133; Conservative 60; Mismatches 145; Indels 102; Gaps 7;

OY 16 RYLVLLNPRGCKKALQFRSHVOPPLAEAEISFTLMTERRNARELVNSELGRMDL 75
DB 216 RYLVLLNPRGCKKALQFRSHVOPPLAEAEISFTLMTERRNARELVNSELGRMDL 275
OY 76 VVMSGDLMEHYVNGIMRPMETAT-OKPLCSLPAGSGMALASLNHYAGYQVYNEDL 134
DB 276 VVMSGDLMEHYVNGIMRPMETAT-OKPLCSLPAGSGMALASLNHYAGYQVYNEDL 134
OY 135 LTRCTLLCRRLSPMLSLHTASGLRFSVLSLNGFTADVDSEKFRRLGEMFTL 194
DB 334 VLGALITYSGSSPMQVYQLOD-RSLTSLTSGMGLSDVDIESRTRMLGQRFY 392
OY 195 GFTPLAALFTYRGRLATPVGRVSGSTPASPVVQGV-----DAHLVPLEEPVSH 248
DB 393 WLYRYLVNFTYNGRISLTLTDEHVSSTHANGYAAORMOSSRCCNHIIMNGPAPY 452
OY 249 -----WT----- 250
DB 453 HSSAEYLQEPADVISLETSTINOSFRSCDMLSGSRSTFYSTISSTYSLADESEFA 512
OY 251 -----VYVDEPFVYVYALHSHSGMFAAP 277
DB 513 GLAASLENRQONTYPASELPDNEPSEDOGLVDEGFVMAHAYOTHTGIDCHAPK 572
OY 278 GRAGVYMLFTYRGVSRAMLRLFLAMEKGRHM-ETPCRYLVNPPVAPRLPKKQK 336
DB 573 AQINDOTITLIRAGISRLHSLSTFLNMSGTHLPESHDDHVKYLPVAPARLEPYDNG 632
OY 337 VFAVDELMVSPAVQGVHP 356
DB 633 IITVDEGEREFGPQAVLP 652

RESULT 15

ID 065419 PRELIMINARY; PRT; 1240 AA.

AC 065419;
DT 01-AUG-1998 (TREMblrel). 07, Created)
DT 01-AUG-1998 (TREMblrel). 07, Last sequence update)
DE 01-OCT-2001 (TREMblrel). 18, Last annotation update)
DE Hypothetical 138.7 kDa protein.
GN F18E5.160 OR AT4G21540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC eurosid 11; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID:3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022603; CA18718.1;
DR EMBL: AL161555; CA81261.1;
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00781; DAGC; 2.
DR ProDom: PD005043; DAG_Kin_cat; 2.
DR SMART: SM00046; DAGC; 2.
DR PROSITE: PS00290; Ig_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ

Query Match 22.48; Score 451; DB 10; Length 1240;
Best Local Similarity 31.98; Pred. No. 1.2e-30;
Matches 122; Conservative 63; Mismatches 127; Indels 70; Gaps 13;

OY 11 LPRCRLVLLNPRGCKKALQFRSHVOPPLAEAEISFTLMTERRNARELVNSELGRMDL 70
DB 374 LPRCRLVLLNPRGCKKALQFRSHVOPPLAEAEISFTLMTERRNARELVNSELGRMDL 433
OY 71 RMDALVMSGDLMEHYVNGIMRPMETAT-OKPLCSLPAGSGMALASLNHYAGYQVYNEDL 130
DB 434 KTDGLVYSGDGLVEYVNGIMRPMETAT-OKPLCSLPAGSGMALASLNHYAGYQVYNEDL 493
OY 131 NEDLITNCTLLCRRLSPMLSLHTASGLRFSVLSLNGFTADVDSEKFRRLGEMFTL 190
DB 494 N-----SATISLIRGHRSDVATI-AQCNTRFFSVLMLANGIADIDISEKFRMGSA 547
OY 191 RETLGTFL-----RLAALFTYRGRLATPV-GRVSGSTPASPVVQGV----- 230
DB 548 RIDFVYCLVDEKFDNYCIAVVKLLALDRIICLRNYGRILFLPAGFGYQGPASGLYGE 607
OY 231 -----OCF-----VDALVPLEEPVSHVTPVDEPFVYVYALHSHSGMFAAP 274
DB 608 PHSVDEKGYQGPETKREDELEMKGPETITM-----LHVPMGSGNTLT 653
OY 275 APMSGCAAGVYMLFTYRGVSRAMLRLFLAMEKGRHM-ETPCRYLVNPPVAPRLPKKQK 336
DB 654 APAAFSGYGLDILVLR-NCPRIVLILSLMROTSSGTHV--ESPFIYIKLTVKVAFLV 710
OY 330 EP-----KDGKGFVAVDGLM 345
DB 711 EFGALVDEPKKEGILDSGEVL 732

Search completed: May 9, 2003, 16:56:56
Job time : 95 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 02:54:38 ; Search time 393 Seconds
(without alignments)
9013.724 Million cell updates/sec

Title: US-09-937-060A-19
Perfect score: 1573
Sequence: 1 gccccacagcgcgcctgcg.....gacctgaaaaaaaaaaaaa 1573

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001D.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1573	100.0	1573	21	AAA75676 DNA encoding a hum
2	1547	98.3	1821	22	AAH15652 Human cDNA sequenc
3	1539.4	97.9	1772	22	AAH16415 Human cDNA sequenc
4	1539	97.8	1719	22	AAD04477 Human sphingosine
5	1511.6	96.1	1600	22	AAD14424 Human sphingosine
6	1479.2	94.0	1533	20	AAV84450 Human secreted pro
7	1479.2	94.0	1533	22	ABA83273 Human secreted pro
8	1430.2	90.9	1562	24	ABL90618 Human poly nucleoti
9	1429.4	90.9	1447	21	AAA50508 Human sphingosine

10	1428.4	90.8	1438	22	AA159336 Human poly nucleoti
11	1168.6	74.3	1205	22	AAC84161 Human sphingosine
12	1165	74.1	1173	24	ABL59532 Human sphingosine
13	1163.2	73.9	1200	22	AA161122 Human poly nucleoti
14	853.4	54.3	1815	21	AA247167 Mouse sphingosine
15	853.2	54.2	1759	22	AAD14425 Mouse consensus sp
16	806.2	51.3	1559	21	AA247168 Mouse sphingosine
17	510.2	32.4	587	22	AAH10443 Human cDNA clone (
18	387.6	24.6	394	24	ABL66446 Lung cancer relate
19	368.8	23.4	675	22	AAH07214 Human cDNA clone (
20	319.6	20.3	604	22	AAH04943 Human cDNA for sph
21	262.8	16.7	2380	22	AAH14817 Human sphingosine
22	262.8	16.7	2380	24	ABL59533 DNA encoding novel
23	262.8	16.7	2422	23	AAH85331 Mouse cDNA for sph
24	246.8	15.7	2698	22	AAH14816 Mouse cDNA for sph
25	240	15.3	240	22	AAD04478 Conserved region D
26	193	12.3	296	22	AAH24477 Human ovarian tumo
27	193	12.3	296	22	AAH83063 Human full-length
28	188.2	12.0	2136	22	AAH94172 Human protein enco
29	181.6	11.5	801	22	AAH99234 Human cDNA 3'-end
30	132.6	8.4	539	22	AAK92477 Human kinase (PKIN
31	116	7.4	1521	22	AAD08644 Human ovarian anti
32	112.8	7.2	1394	24	ABQ54324 DNA encoding novel
33	112.8	7.2	1578	23	AAH85330 Human nervous syst
34	112.8	7.2	11096	22	ABA18278 Drosophila melanog
35	110.6	7.0	2830	24	ABL02327 Human ORFX polynuc
36	105.8	6.7	297	24	ABN22070 Drosophila melanog
37	91.4	5.8	2020	23	ABL03297 Drosophila melanog
38	91.4	5.8	4020	23	ABL03296 Drosophila melanog
39	83	5.3	5686	23	ABN1630 Human spliced tran
40	60	3.8	60	24	ABN1630 Human spliced tran
41	57	3.6	426	24	ABL51829 A1a-Pro-A1a-pro co
42	52.2	3.3	2744	16	AAO98470 M.SPI-containing p
43	51.2	3.3	547	21	AAH76592 Human ORFX ORF2147
44	51.2	3.3	4231	21	AAH76592 Human sphingosine
45	51.2	3.3	4413	24	ABL40828 Human sphingosine

ALIGNMENTS

RESULT 1	
AA75676	AAA75676 standard: DNA: 1573 BP.
AC	AAA75676;
XX	
XX	
DT	22-JAN-2001 (first entry)
XX	
DE	DNA encoding a human regulator of intracellular phosphorylation.
XX	
KW	Human; intracellular phosphorylation regulator; HRP; stroke; myeloma;
KW	neurological disorder; Parkinson's disease; demyelinating disease;
KW	meningitis; developmental disorder; neuromuscular disorder; cancer;
KW	myasthenia gravis; cell proliferative disorder; actinic keratosis;
KW	arteriosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;
KW	autoimmune disorder; inflammatory disorder; Addison's disease;
KW	acquired immunodeficiency disease; allergy; diabetes mellitus;
KW	rheumatoid arthritis; microbial infection; trauma; ss.
OS	Homo sapiens.
XX	
XX	
FB	Key
FT	Location/Qualifiers
FT	130..1284
FT	CDS
FT	/*tag= a
FT	/product= "regulator of intracellular phosphorylation"
XX	
XX	W020005332-A2.
XX	
PD	21-SEP-2000.
XX	
XX	
PF	17-MAR-2000; 2000MO-US07277.
XX	

PR 18-MAR-1999; 99US-0125593.
PR 20-MAY-1999; 99US-0135049.
PR 09-JUL-1999; 99US-0143188.
XX
PA (INCY-) INCYTE PHARM INC.
PI Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
PI Lu DM, Au-Young J;
XX
XX WPI: 2000-602121/57.
DR P-PSDB; AABI8659.
XX
XX Novel human intracellular phosphorylation regulator polypeptides and
PT polynucleotides for diagnosis, prevention and treatment of
PT neurological, cell proliferative and autoimmune/inflammatory disorders
PT
PT
DS Claim 4; Page 89; 96pp: English.

The present sequence encodes a human regulator of intracellular phosphorylation (HRIP). HRIP is useful for screening agonists and antagonists of HRIP polypeptide. HRIP and its agonist or antagonist are useful for treating a disease or condition associated with decreased or increased expression of functional HRIP. Diseases treated or diagnosed include neurological disorders such as stroke, Parkinson's disease, demyelinating diseases, bacterial and viral meningitis and other developmental disorders of the central nervous system, neuromuscular disorders, myasthenia gravis, cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer including leukemia, melanoma, myeloma and cancer of the adrenal gland, bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/inflammatory disorder such as Addison's disease, acquired immunodeficiency diseases, allergies, bronchitis, diabetes mellitus, rheumatoid arthritis, microbial infection and trauma.

SQ Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other;

Query Match	100.0%;	Score 1573;	DB 21;	Length 1573;
-------------	---------	-------------	--------	--------------

Matches 1573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGCCCAACAGCGGGCCCTGTGGAGCGCCCTGTGGGCAAGCAACGATTAAGAGCTGAAGCGAG	60
Db	1	GGCCCAACAGCGGGCCCTGTGGAGCGCCCTGTGGGCAAGCAACGATTAAGAGCTGAAGCGAG	60
Qy	61	ACCGCGCGCCACCGGGCAAGCGCCCAACAGCGCCACCGGAGCCCTTGGCAGCGGGAGCGCGG	120
Db	61	ACCGCGCGCCACCGGGCAAGCGCCCAACAGCGCCACCGGAGCCCTTGGCAGCGGGAGCGCGG	120
Qy	121	GTGAGAGTTATGAGATCAACGGGGGGGGCCCGGGGGGTGTCTCCCGCGGCTCTGCCCGCTG	180
Db	121	GTGAGAGTTATGAGATCAACGGGGGGGGCCCGGGGGGTGTCTCCCGCGGCTCTGCCCGCTG	180
Qy	181	CTGGAGCTGCTGAACCCGGCGCGCGGCAAGGGCAAGGGCTTGGAGCTCTTCGCGAGGTAC	240
Db	181	CTGGAGCTGCTGAACCCGGCGCGCGGCAAGGGCAAGGGCTTGGAGCTCTTCGCGAGGTAC	240
Qy	241	GTGCAGCCCCCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTAGCGCGCG	300
Db	241	GTGCAGCCCCCTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTAGCGCGCG	300
Qy	301	AACCAAGCGCGGAGCTGTGTGCGGTCCGAGAGACTGGCCCGCTGGACGCTGTGGTGTGTC	360
Db	301	AACCAAGCGCGGAGCTGTGTGCGGTCCGAGAGACTGGCCCGCTGGACGCTGTGGTGTGTC	360
Qy	361	ATGCTGAGACCGGCTGATGTAGAGAGTGTGTAAACGGGCTCAATGAGAGCGGCGCTGACACG	420
Db	361	ATGCTGAGACCGGCTGATGTAGAGAGTGTGTAAACGGGCTCAATGAGAGCGGCGCTGACACG	420
Qy	421	GAGACCGGCATCCAGAGCCCGCTGTGTAGCTTCCAGCAGGCTCTGTGCAACGCGCTGGCA	480
Db	421	GAGACCGGCATCCAGAGCCCGCTGTGTAGCTTCCAGCAGGCTCTGTGCAACGCGCTGGCA	480

QY	481	GCCTCCCTGAACCACTTATGCTGCTATGAGCGAGGTGACATGATGAAGCCCTGACAC	540
Db	481	GCCTCCCTGAACCACTTATGCTGCTATGAGCGAGGTGACATGATGAAGCCCTGACAC	540
QY	541	TGCACGCTATTGCTGTGCTCCGCGGCTGCTACCCATGAACCTGCTCTCTCCACAG	600
Db	541	TGCACGCTATTGCTGTGCTCCGCGGCTGCTACCCATGAACCTGCTCTCTCCACAG	600
QY	601	GGTTGGGGGGCTGCGCCCTTCTCTGTGTGTGACACCTGGGCTGGGCTTCAATGCTATG	660
Db	601	GGTTGGGGGGCTGCGCCCTTCTCTGTGTGTGACACCTGGGCTGGGCTTCAATGCTATG	660
QY	661	GACCTAGAGATGAGAAAGTATCGGCTCTGGGGGAGATGGCTTCACTCTGGGACCTTC	720
Db	661	GACCTAGAGATGAGAAAGTATCGGCTCTGGGGGAGATGGCTTCACTCTGGGACCTTC	720
QY	721	CTGGCTCTGGGAGCCCTTGGCCACTTACGGGGCGGACTGGCTACCTCCCTGTAGGAAGA	780
Db	721	CTGGCTCTGGGAGCCCTTGGCCACTTACGGGGCGGACTGGCTACCTCCCTGTAGGAAGA	780
QY	781	GTGGGTTTCAAGACACCTGGCTCCCGGTGTGTGTCAGCAGGGGCCCGGTAGATGCACAC	840
Db	781	GTGGGTTTCAAGACACCTGGCTCCCGGTGTGTGTCAGCAGGGGCCCGGTAGATGCACAC	840
QY	841	CTTGTGACACATGAGAGAGCCAGTGGCCCTTCACTGTGACAGTGGTGGCCGACAGACATTT	900
Db	841	CTTGTGACACATGAGAGAGCCAGTGGCCCTTCACTGTGACAGTGGTGGCCGACAGACATTT	900
QY	901	GTGCTAGTCCCTGGGACATGCTGCACCTGCACCTGGGCGAGTGAATGTTTCTGACCCATG	960
Db	901	GTGCTAGTCCCTGGGACATGCTGCACCTGCACCTGGGCGAGTGAATGTTTCTGACCCATG	960
QY	961	GGCCGCTGTGAGTGGCGCATGATCATGCTGCTTCACTGCGCGCGGAGGCTGTCTGCC	1020
Db	961	GGCCGCTGTGAGTGGCGCATGATCATGCTGCTTCACTGCGCGCGGAGGCTGTCTGCC	1020
QY	1021	ATGCTGCTGCGCCCTCTTCTGGGCGATGGAAGGGCAGGCACTATGAGTATGAATGCCCC	1080
Db	1021	ATGCTGCTGCGCCCTCTTCTGGGCGATGGAAGGGCAGGCACTATGAGTATGAATGCCCC	1080
QY	1081	TACTTGTATATGTGCGCGGTGGCGCTTCCGCTTGGAGGCCAAAGGTGGGAAAGGTGTG	1140
Db	1081	TACTTGTATATGTGCGCGGTGGCGCTTCCGCTTGGAGGCCAAAGGTGGGAAAGGTGTG	1140
QY	1141	TTTGGAGTGAATGGGGAATGATGGTTAGCAGAGCCGTGCAAGGGCCAGCTGACCCCAAC	1200
Db	1141	TTTGGAGTGAATGGGGAATGATGGTTAGCAGAGCCGTGCAAGGGCCAGCTGACCCCAAC	1200
QY	1201	TACTTGTGATGCTCAGCGGTGTGCTGAGAGCCCGCCAGACTGTGAAAGGCCCAAGCATG	1260
Db	1201	TACTTGTGATGCTCAGCGGTGTGCTGAGAGCCCGCCAGACTGTGAAAGGCCCAAGCATG	1260
QY	1261	CCACGCGCAGAAAGACCCTTATGACCCCTGGGCGCGCTGTGCTTAGTGTCTACTTGA	1320
Db	1261	CCACGCGCAGAAAGACCCTTATGACCCCTGGGCGCGCTGTGCTTAGTGTCTACTTGA	1320
QY	1321	GGACCCCTTCCCTCCTTCCCTAGGGCTGTGCAAGGCTGTCCACAGCTCTCTGTGGGGGTGAGC	1380
Db	1321	GGACCCCTTCCCTCCTTCCCTAGGGCTGTGCAAGGCTGTCCACAGCTCTCTGTGGGGGTGAGC	1380
QY	1381	AGACTTCCTGTGACAAAGGTGGAAGAGTGTGAGAGCTTTTGGGGGGACAGGCCAGAT	1440
Db	1381	AGACTTCCTGTGACAAAGGTGGAAGAGTGTGAGAGCTTTTGGGGGGACAGGCCAGAT	1440
QY	1441	GAAAGTCCGGGTCAGAGGCCCAAGCTGGCTGGGGCCACGTGCTATGTAAGGGCTTCTAGT	1500
Db	1441	GAAAGTCCGGGTCAGAGGCCCAAGCTGGCTGGGGCCACGTGCTATGTAAGGGCTTCTAGT	1500
QY	1501	TTGTGTTGAGAGCCCAAGCCCAAGACCAAAATCCAAATGAAGTACATTGCCAGCTGAA	1560
Db	1501	TTGTGTTGAGAGCCCAAGCCCAAGACCAAAATCCAAATGAAGTACATTGCCAGCTGAA	1560
QY	1561	AAAAAAAAAAAAA 1573	

QY	481	GCCTCCCTGAACCACTTATGCTGCTATGAGCGAGGTGACATGATGAAGCCCTGACAC	540
Db	481	GCCTCCCTGAACCACTTATGCTGCTATGAGCGAGGTGACATGATGAAGCCCTGACAC	540
QY	541	TGCACGCTATTGCTGTGCTCCGCGGCTGCTACCCATGAACCTGCTCTCTCCACAG	600
Db	541	TGCACGCTATTGCTGTGCTCCGCGGCTGCTACCCATGAACCTGCTCTCTCCACAG	600
QY	601	GGTTGGGGGGCTGCGCCCTTCTCTGTGTGCTACCCCTGGGCTGGGCTTCAATGCTATG	660
Db	601	GGTTGGGGGGCTGCGCCCTTCTCTGTGTGCTACCCCTGGGCTGGGCTTCAATGCTATG	660
QY	661	GACCTAGAGATGAGAAAGTATCGCCCTCTGGGGGAGATGGCTTCACTCTGGGACCTTC	720
Db	661	GACCTAGAGATGAGAAAGTATCGCCCTCTGGGGGAGATGGCTTCACTCTGGGACCTTC	720
QY	721	CTGGCTCTGGCAGCCCTTGGCCACTTACGGGGCGAGACTGGCTACCTCCCTGTAGGAAGA	780
Db	721	CTGGCTCTGGCAGCCCTTGGCCACTTACGGGGCGAGACTGGCTACCTCCCTGTAGGAAGA	780
QY	781	GTGGGTTTCAAGACACCTGGCTCCCGGTGTGTGTCACGAGGGCCCGGTAGATGCACAC	840
Db	781	GTGGGTTTCAAGACACCTGGCTCCCGGTGTGTGTCACGAGGGCCCGGTAGATGCACAC	840
QY	841	CTTGTGACACATGAGAGAGCCAGTGGCCCTTCACTGTGACAGTGGTGGCCGACAGACATTT	900
Db	841	CTTGTGACACATGAGAGAGCCAGTGGCCCTTCACTGTGACAGTGGTGGCCGACAGACATTT	900
QY	901	GTGCTAGTCCCTGGGACATGCTGCACCTGCACCTGGGCGAGTGAATGTTTCTGACCCATG	960
Db	901	GTGCTAGTCCCTGGGACATGCTGCACCTGCACCTGGGCGAGTGAATGTTTCTGACCCATG	960
QY	961	GGCCGCTGTGAGTGGCGCATGATCATGCTCTTCACTGCGCGCGGAGGCTGTCTGCC	1020
Db	961	GGCCGCTGTGAGTGGCGCATGATCATGCTCTTCACTGCGCGCGGAGGCTGTCTGCC	1020
QY	1021	ATGCTGCTGCGCCCTCTTCTGGGCGATGGAAGGGCGAGGCTATGAGTATGAATGCCCC	1080
Db	1021	ATGCTGCTGCGCCCTCTTCTGGGCGATGGAAGGGCGAGGCTATGAGTATGAATGCCCC	1080
QY	1081	TACTTGTGATATGTGCGCGGTGGCGCTTCCGCTTGGAGCCCAAGGATGGGAAGGTGTG	1140
Db	1081	TACTTGTGATATGTGCGCGGTGGCGCTTCCGCTTGGAGCCCAAGGATGGGAAGGTGTG	1140
QY	1141	TTTGGAGTGAATGGGGAATGATGGTATAGCAGGCGCTGCAAGGGCCAGGCTGACCCCAAC	1200
Db	1141	TTTGGAGTGAATGGGGAATGATGGTATAGCAGGCGCTGCAAGGGCCAGGCTGACCCCAAC	1200
QY	1201	TACTTGTGATATGTGCGCGGTGGCGCTTCCGCTTGGAGCCCAAGGATGGGAAGGTGTG	1260
Db	1201	TACTTGTGATATGTGCGCGGTGGCGCTTCCGCTTGGAGCCCAAGGATGGGAAGGTGTG	1260
QY	1261	CCACGCGCAGAGAGCCCTTATGACCCCTGGGCGCGCTGTGCTTAGTGTCTACTTGCAT	1320
Db	1261	CCACGCGCAGAGAGCCCTTATGACCCCTGGGCGCGCTGTGCTTAGTGTCTACTTGCAT	1320
QY	1321	GGACCCCTTCCCTCCTTCCCTTAGGGCTGTCAGAGGCTGTCCACAGCTCTCTGTGGGGGTGAGC	1380
Db	1321	GGACCCCTTCCCTCCTTCCCTTAGGGCTGTCAGAGGCTGTCCACAGCTCTCTGTGGGGGTGAGC	1380
QY	1381	AGACTTCCTGTGACAAAGGTGGAAGAGTGTGAGAGCTTATGTTGGGGGGACAGCCAGAT	1440
Db	1381	AGACTTCCTGTGACAAAGGTGGAAGAGTGTGAGAGCTTATGTTGGGGGGACAGCCAGAT	1440
QY	1441	GAAAGTCCGGGCTGAGAGGCCACAGTGGCTGGGGCCACGCTGCTATGTAAGGCTTCTAGT	1500
Db	1441	GAAAGTCCGGGCTGAGAGGCCACAGTGGCTGGGGCCACGCTGCTATGTAAGGCTTCTAGT	1500
QY	1501	TTGTTCTGTAGAGCCGCCACGCCACAGACCAATTCACAAATGAAGTACATTGCCAGCTGAA	1560
Db	1501	TTGTTCTGTAGAGCCGCCACGCCACAGACCAATTCACAAATGAAGTACATTGCCAGCTGAA	1560
QY	1561	AAAAAAAAAAAAA 1573	

DB 1561 AAAAAAAAAA 1573

RESULT 2

AAH15652
ID AAH15652 standard; cDNA: 1821 BP.

AAH15652;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:13996.

Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI: 2001-318749/34.

Claim 8: SEQ ID 13996; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95883 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1821 BP; 301 A; 569 C; 614 G; 337 T; 0 other;

Query Match 98.3%; Score 1547; DB 22; Length 1821;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCCCCACAGCCGCGCTGCGAGCCCGCTGGGAGACCGATTAAGAGCTAAGGACGAG 60
DB 263 GCCCCACAGCCGCGCTGCGAGCCCGCTGGGAGACCGATTAAGAGCTAAGGACGAG 322
QY 61 AGCCCCCGCCACAGGCGGACGG -CCCCACAGCGCCAGAGGAGCCCTGGACAGCGGAGCGCG 119
DB 323 AGCCGCGCGCCACAGGCGGAGCGCCCGCCACAGCGCGGAGGAGCCCTGGACAGCGGAGCGCG 382
QY 120 GGTGAGGTTATGATATCCAGCGGCGGCGCCCGGCGGCGTGTCCCGCGCGCCGCGCGCT 179
DB 383 GGTGAGGTTATGATATCCAGCGGCGGCGCCCGGCGGCGTGTCCCGCGCGCCGCGCGCT 442
QY 180 GCTGTGCTGCTGAACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 239
DB 443 GCTGTGCTGCTGAACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
QY 240 CGTGAGGCGCGCTTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTCTACGAGCGCGG 299
DB 503 CGTGAGGCGCGCTTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTCTACGAGCGCGG 562
QY 300 GAACACAGCGCGGCGGAGCTGTGCGGCGGAGAGAGCTGGGCGGCGGCGGCGGCGGCGGCGG 359
DB 563 GAACACAGCGCGGCGGAGCTGTGCGGCGGAGAGAGCTGGGCGGCGGCGGCGGCGGCGGCGG 622
QY 360 CATGCTTGAGAGACGGGCTGATGACAGAGGTGTGAACGGGCTCATAGAGCGGCGGCTGAC 419
DB 623 CATGCTTGAGAGACGGGCTGATGACAGAGGTGTGAACGGGCTCATAGAGCGGCGGCTGAC 682
QY 420 GAGAGCGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479
DB 683 GAGAGCGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
QY 480 AGCTTCCTTGAACCATTAATGCTGCGGCTGATGAGAGAGGCTGACCAATGAAGAGCTTCGAC 539
DB 743 AGCTTCCTTGAACCATTAATGCTGCGGCTGATGAGAGAGGCTGACCAATGAAGAGCTTCGAC 802
QY 540 CTGACAGCTATTGCTGTGCGCGCGGCTGTGACCCATGAACCTGCTCTCTGACAC 599
DB 803 CTGACAGCTATTGCTGTGCGCGCGGCTGTGACCCATGAACCTGCTCTCTGACAC 862
QY 600 GCGCTTCGGGCGGCGGCGGCTGTGCTGTGCTGACAGCTGCGGCGGCGGCGGCGGCGGCGG 659
DB 863 GCGCTTCGGGCGGCGGCGGCTGTGCTGTGCTGACAGCTGCGGCGGCGGCGGCGGCGGCGG 922
QY 660 GAGCCTAGAGAGTGAAGATATGCGGCTGCGGCGGAGATGCGGCTTACCTGGGACCT 719
DB 923 GAGCCTAGAGAGTGAAGATATGCGGCTGCGGCGGAGATGCGGCTTACCTGGGACCT 982
QY 720 CCGGCTGCTGGGAGCGCTGCGGAGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779
DB 983 CCGGCTGCTGGGAGCGCTGCGGAGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1042
QY 780 AGTGGCTTCCAGACACCTGCTCCCGGCTGTGTGCTGACAGAGGCGGCGGCTGATGAC 839
DB 1043 AGTGGCTTCCAGACACCTGCTCCCGGCTGTGTGCTGACAGAGGCGGCGGCTGATGAC 1102
QY 840 CCGTGGGCGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 899
DB 1103 CCGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1162
QY 900 TGTGCTAGTCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 959
DB 1163 TGTGCTAGTCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 1222
QY 960 GGGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 1019
DB 1223 GGGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 1282
QY 1020 CATGCTGCGGCGGCTGCTGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1079
DB 1283 CATGCTGCGGCGGCTGCTGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1342
QY 1080 CTACTTGATATATGTCGCCGTGTGCTGCGCTTGGAGCGGCGGAGGAGGAGGAGGAGGAGG 1139

Db 1343 CTACTTGTATATGTCCTCCGTCCTCCCTTCCGAGCCCAAGGATGGAGAGTGT 1402
 Oy 1140 GTTTCGATGATGAGGAAATGATGTTACGAGCCCTCCAGGCGCCAGTCCACCCAAA 1199
 Db 1403 GTTTCGATGATGAGGAAATGATGTTACGAGCCCTCCAGGCGCCAGTCCACCCAAA 1462
 Oy 1200 CTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 Db 1463 CTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1522
 Oy 1260 GCCACCCGCAAGAGAGCCCTTATGACCCCTGCGCGCGCTGCTTACTTCTACTTGC 1319
 Db 1533 GCCACCCGCAAGAGAGCCCTTATGACCCCTGCGCGCGCTGCTTACTTCTACTTGC 1592
 Oy 1320 AGGACCTTCT 1379
 Db 1583 AGGACCTTCT 1642
 Oy 1380 GAGACT 1439
 Db 1643 GAGACT 1702
 Oy 1440 TGAAGTCT 1499
 Db 1703 TGAAGTCT 1762
 Oy 1500 TTTGTTCT 1558
 Db 1763 TTTGTTCT 1821

RESULT 3
 AAH16415
 ID AAH16415 standard; cDNA: 1772 BP.
 AC AAH16415:
 DT 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:15393.
 DE Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 KW Homo sapiens.
 OS Homo sapiens.
 PN EPI074617-A2.
 XX 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8: SEQ ID 15393: 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides, and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SO Sequence 1772 BP: 285 A: 566 C: 585 G: 336 T: 0 other:
 Query Match 97.9%; Score 1539.4; DB 22; Length 1772;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1551; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 1 GCCCCACAGCCCGCCCTGCGAGCCCGCTGGCCAGCAGCCGATTAAGGAGCGAAGCCAGG 60
 Db 220 GCCCCACAGCCCGCCCTGCGAGCCCGCTGGCCAGCAGCCGATTAAGGAGCGAAGCCAGG 279
 Oy 61 AGCCCGCCCGCCAGCCGAGCG-CCCCACAGCCGAGCCGAGCCCTGGCAGCGGAGCCGCG 119
 Db 280 AGCCCGCCCGCCAGCCGAGCGCCGCGCCAGCCGAGCCGAGCCCTGGCAGCGGAGCCGCG 339
 Oy 120 GGTTCGAGTTATGATTCAGCGGGGCGCCCGGGGCTGTCTCCGCGGCTCCGCGCT 179
 Db 340 GGTTCGAGTTATGATTCAGCGGGGCGCCCGGGGCTGTCTCCGCGGCTCCGCGCT 399
 Oy 180 GCTGTCGTCGTGAACCCCGCGCGGCGCAAGGGCTTGGACGCTCTCCGCGAGTCA 239
 Db 400 GCTGTCGTCGTGAACCCCGCGCGGCGCAAGGGCTTGGACGCTCTCCGCGAGTCA 459
 Oy 240 CGTGACAGCCCTTTTGGCTGAGCGCTGAATCTCTTACGCTGATGCTCAGTGAAGCCGCG 299
 Db 460 CGTGACAGCCCTTTTGGCTGAGCGCTGAATCTCTTACGCTGATGCTCAGTGAAGCCGCG 519
 Oy 300 GAACCAAGCGCGGAGCTGTGCGGTGAGAGAGCTGGGCGCTGGAGCGCTGTGGT 359
 Db 520 GAACCAAGCGCGGAGCTGTGCGGTGAGAGAGCTGGGCGCTGGAGCGCTGTGGT 579
 Oy 360 CATGTCGTCGAGCGCGCTGATGCAAGAGCTGTGCAAGGGCTTGGACGCTCTCCGCGAGTCA 419
 Db 580 CATGTCGTCGAGCGCGCTGATGCAAGAGCTGTGCAAGGGCTTGGACGCTCTCCGCGAGTCA 639
 Oy 420 GGAGACCGCCATCCAGAGCCCTGTAGCTCCAGCAGGCTGTGGCAAGCGGCTGGC 479
 Db 640 GGAGACCGCCATCCAGAGCCCTGTAGCTCCAGCAGGCTGTGGCAAGCGGCTGGC 699
 Oy 480 AGCTTCCTGAACCAATTAATGCTGATGAGAGCTGATGACCATTAAGAGCTCTGACCA 539
 Db 700 AGCTTCCTGAACCAATTAATGCTGATGAGAGCTGATGACCATTAAGAGCTCTGACCA 759
 Oy 540 CTGACGCTATTTGTCGTCGCGCGGCTGTCACCATTAAGAGCTCTGTCGACAC 599
 Db 760 CTGACGCTATTTGTCGTCGCGCGGCTGTCACCATTAAGAGCTCTGTCGACAC 819
 Oy 600 GGGTTGGGGGTCGCGCTCTTCTGTGCTACGCTGAGCGCTGGGCTTCATTTGATGT 659
 Db 820 GGGTTGGGGGTCGCGCTCTTCTGTGCTACGCTGAGCGCTGGGCTTCATTTGATGT 879
 Oy 660 GAGCTAGAGATGAGAGTATCGGCTCTGCGGGGAGATGCCCTTCACTCTGCGCACCTT 719


```

|||||
Db 880 GGACCTAGAGAGTAGAGATGCGCGCTGTGGGAGAGATGCGCTTACACTGTGGGACCTT 939
Oy 720 CCTCGTGTGACAGCCCTGCGCACCTACCGGGCGGAGCTGACCTACCTCCCTGAAG 779
Db 940 CCTCGTGTGACAGCCCTGCGCACCTACCGGGCGGAGCTGACCTACCTCCCTGAAG 999
Oy 780 AGTGGTTTCCAAAGACACCTGCGCTCCCGGTTGTGTCCAGAGAGGCGCGTAGATGACA 839
Db 1000 AGTGGTTTCCAAAGACACCTGCGCTCCCGGTTGTGTCCAGAGAGGCGCGTAGATGACA 1059
Oy 840 CCTGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
Db 1060 CCTGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1119
Oy 900 TGTGCTACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
Db 1120 TGTGCTACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
Oy 960 GGGCGGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
Db 1180 GGGCGGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
Oy 1020 CATGCTGTGCGGCTCTTCTGCGCAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
Db 1240 CATGCTGTGCGGCTCTTCTGCGCAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
Oy 1080 CTACTGTGTATATGCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 1139
Db 1300 CTACTGTGTATATGCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 1359
Oy 1140 GTTTCAGATGATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
Db 1360 GTTTCAGATGATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
Oy 1200 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
Db 1420 CTACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Oy 1260 GCGACGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
Db 1480 GCGACGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
Oy 1320 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1379
Db 1540 AGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1599
Oy 1380 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439
Db 1600 GAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1659
Oy 1440 TGAAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
Db 1660 TGAAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1719
Oy 1500 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1552
Db 1720 TTTGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772

```

RESULT 4
AAD04477 standard; cDNA; 1719 BP.

XX AAD04477;

XX 04-JUL-2001 (first entry)

DE Human sphingosine kinase type 1 (hsk1) cDNA.

KW Human: sphingosine kinase type 1; sk1; chromosome 17q25.2;
KW sphingosine-1-phosphate; SIP; drug screening; therapy; haemostasis;
KW thrombosis; allergic reaction; proliferative disease; cancer;

```

KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;  
KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;  
KW autoimmune disease; inflammatory disease; multiple sclerosis;  
KW T helper-1 related disease; chronic obstructive pulmonary disease;  
KW asthma; myocardial infarction; neurodegenerative disorder;  
KW wound healing; embryogenesis; anticoagulant; cerebroprotective;  
KW neuroprotective; antiproliferative; antiarthritic; cyostatic; cardiac;  
KW veterinary; ss.  
XX  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 270..1424  
FT     /tag=a  
FT     /product="Human sphingosine kinase type 1 (hsk1)"  
FT     /note="CDS is specifically claimed in claim 2 and  
FT     shown as SEQ ID NO 2"  
FT     1675..1681  
FT     /tag=b  
FT     264..273  
FT     misc_feature  
FT     /tag=c  
FT     /note="Translational initiator AUG is in a partial  
FT     kozak consensus"  
FT  
XX  
XX MO200131029-A2.  
XX  
PD 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-EP09498.  
XX  
XX 28-OCT-1999; 99US-0162307.  
XX 28-OCT-1999; 99US-0162307.  
XX 07-FEB-2000; 2000US-0180525.  
XX  
XX (WARN ) WARNER LAMBERT CO.  
XX  
XX Allen J, Goslink M, Melendez AJ, Takacs L;  
XX  
XX WPI: 2001-300510/31.  
XX  
XX P-PSDB: AAE00924.  
XX  
XX New human sphingosine kinase type I gene for screening drug candidates  
XX particularly inhibitors used for preventing or treating e.g.  
XX atherosclerosis, thrombosis, asthma and diabetes  
XX  
XX  
XX  
XX Claim 2: Fig 1; 91pp: English.  
XX  
XX  
XX The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.  
XX The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the  
XX substrate sphingosine to sphingosine-1-phosphate (SIP). The sk1 gene  
XX and encoded polypeptide are applicable in screening drug candidates  
XX particularly inhibitors for preventing or treating disorders such as  
XX haemostasis, thrombosis, allergic reactions, proliferative diseases  
XX including cancer, haematopoietic disorders such as leukaemia,  
XX cardiovascular diseases such as stroke, atherosclerosis and coronary  
XX artery disease, dyslipidaemia, diabetes including type I and type II  
XX diabetes, autoimmune and inflammatory diseases such as multiple  
XX sclerosis, T helper-1 related diseases, chronic obstructive pulmonary  
XX disease, asthma, myocardial infarction, neurodegenerative disorders,  
XX natural wound healing processes and embryogenesis.  
XX  
XX Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other:  
SQ  


```

Query Match 97.8%; Score 1539; DB 22; Length 1719;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1548; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Oy 11 GGGCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 70
Db 151 GGGCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 210
Oy 71 ACGGCGAGCGCCACAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 130
Db 211 ACGGCGAGCGCCACAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270

```


XX The present invention relates to sphingosine kinase (Sphk) polypeptides
 CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-
 CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischemia and cell proliferative disease
 CC or disorder associated with vascular diseases. Sphk gene is used in gene
 CC therapy and antisense-therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth,
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is human sphingosine kinase
 CC (Sphk) cDNA.

XX Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other:

Query Match 96.1%; Score 1511.6; DB 22; Length 1600;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 1334; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

1 GCCCCACACGCGCCCTGGACGCGCCCTGGGACACACGATAGAGCTGAAGGACG 60
 46 GCCCCACACGCGCCCTGGACGCGCCCTGGGACACACGATAGAGCTGAAGGACG 105
 61 AGCGCGCGCCACGCGGACGCG-CCCCACACGCGCGGAGGAGCCCTGGACGCGCGCG 119
 106 AGCGCGCGCGCGGACGCGCGCCCGCCACAGCGCGGAGGAGCCCTGGACGCGCGCG 165
 120 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGCGGT 179
 166 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGCGGT 225
 180 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 229
 226 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 285
 240 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 299
 286 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 345
 300 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 359
 346 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 405
 360 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 419
 406 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 465
 420 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 479
 466 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 525
 480 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 539
 526 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 585
 540 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 599
 586 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 645
 600 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 659
 646 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 705
 660 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 719
 706 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 765
 720 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 779
 766 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 825
 780 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 839

DB 826 AGTGGGTTTCACAGACCTGCTCCCGCGGTGTGTGTCACAGGCGCGGTATGTCACA 885
 840 CCTTGGCCACCTGAGAGGACCTGCGCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 899
 886 CCTTGGCCACCTGAGAGGACCTGCGCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 945
 900 TGTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
 946 TGTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
 960 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1019
 1006 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1065
 1020 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1079
 1066 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1125
 1080 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1139
 1126 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1185
 1140 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1199
 1186 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1245
 1200 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1259
 1246 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1305
 1260 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1319
 1306 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1365
 1320 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1379
 1366 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1425
 1380 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1439
 1426 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1485
 1440 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1499
 1486 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1545
 1500 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1553
 1546 GGTCCAGGTTATGATTCACAGCGGCGCGCCCGCGGCGGTCTCCGCGCGCGGTCTCA 1599

RESULT 6
 AA84490
 ID AA84490 standard; DNA: 1533 BP.
 XX
 AC AA84490:
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Human secreted protein gene 80 clone HNF454.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostatic; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX

CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
 CC in gene therapy and vaccine production. (I) and (II) can be used in the
 CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gnucler's disease), hyperproliferative disorders (e.g. scleritis, angiodermic
 CC Chagas's cardiomyopathy and coronary arteriosclerosis), angiodermic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. ABA3185 to
 CC ABA3193 and ABA30300 represent sequences used in the exemplification of
 CC the present invention.

XX Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other:

Query Match 94.08; Score 1479.2; DB 22; Length 1533;

Best Local Similarity 98.88; Pred. No. 0;

Matches 1492; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 65 GCCGCCACGGGCGGCGCCACAGCGCCAGGACCCCTGCGAGCGGGAGCGCGGGCTCG 124
 DB 8 GCGGACACGGGCGGCGCCATAGCGCCAGGACCCCTGCGAGCGGGAGCGCGGGCTCG 67
 QY 125 AGCTTATGATCCAGCGGGGCGCGCGCGCGCGCTGCTCCGCGGCGCTGCGCGCTG 184
 DB 68 AGGTTATGATCCAGCGGGGCGCGCGCGCGCGCTGCTCCGCGGCGCTGCGCGCTG 127
 QY 185 TGGTGTGAACCGCGCGCGCGCGCGCGCGCGCGCTGCGAGCTTCCGAGTACAGTGC 244
 DB 128 TGGTGTGAACCGCGCGCGCGCGCGCGCGCGCGCTGCGAGCTTCCGAGTACAGTGC 187
 QY 245 AGCCCTTTGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 304
 DB 188 AGCCCTTTGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 247
 QY 305 AGCCGCGGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 364
 DB 248 AGCCGCGGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 307
 QY 365 CTGAGAGCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 423
 DB 308 YTGAGAGCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 367
 QY 424 ACCGCGATCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 483
 DB 368 ACCGCGATCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 427
 QY 484 TCCCTGAACCATTTATGCTGCTATGAGCAGGTACCAATGAAGACCTCTGACCACTGC 543
 DB 428 TCCCTGAACCATTTATGCTGCTATGAGCAGGTACCAATGAAGACCTCTGACCACTGC 487
 QY 544 AGCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 603
 DB 488 AGCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 547
 QY 604 TCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 663
 DB 548 TCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 607
 QY 664 CTAGAGAGTGAAGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 723
 DB 608 CTAGAGAGTGAAGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 667
 QY 724 CGTCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 783
 DB 668 CGTCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 727
 QY 784 GGTTCAGAGACCTGCTCCCGCGTGTGTGTCAGAGAGGCGCGGTAGATGACACACTT 843
 DB 728 GGTTCAGAGACCTGCTCCCGCGTGTGTGTCAGAGAGGCGCGGTAGATGACACACTT 787
 QY 844 GTGCCACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 903

DB 788 GTGCCACTGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 847
 QY 904 CTAGTCTGCGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 963
 DB 848 CTAGTCTGCGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 907
 QY 964 CCTGTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1023
 DB 908 CCTGTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 967
 QY 1024 CTGTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1083
 DB 968 CTGTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1027
 QY 1084 TTGTATATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1143
 DB 1028 TTGTATATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1087
 QY 1144 GCAAGTATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1203
 DB 1088 GCAAGTATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1147
 QY 1204 TTCTGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1263
 DB 1148 TTCTGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1207
 QY 1264 CCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1323
 DB 1208 CCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1267
 QY 1324 CCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1383
 DB 1268 CCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1327
 QY 1384 CTCTCTGCGAGCGGTCGAGAGGTCGAGAGGTCGAGAGGTCGAGAGGTCGAGAGGTC 1443
 DB 1328 CTCTCTGCGAGCGGTCGAGAGGTCGAGAGGTCGAGAGGTCGAGAGGTCGAGAGGTC 1387
 QY 1444 GTCTGAGTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1503
 DB 1388 GTCTGAGTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1447
 QY 1504 TTCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1563
 DB 1448 TTCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGTACAGTGC 1507
 QY 1564 AAAAAAAAAA 1573
 DB 1508 AAAAAAAAAA 1517

RESULT 8

ABL90618/C

ID ABL90618 standard; cDNA; 1562 BP.

AC ABL90618;

DT 24-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 1180.

XX

XX Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;

XX antitumor; hepatotropic; antidiabetic; antileukemia; antifungal;

XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX neurological disease; infection; human; secreted protein; gene; ss.

OS

XX Homo sapiens.

PN

XX W0200190304-A2.

XX

XX 29-NOV-2001.


```
QY 1085 TGGTATATGCCCCGTCGTCGCTTCCCTTGAGCCCAAGATGGAAAGGTGTGTTG 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 TGGTATAGTGGCCCGTGGTGGCTTCCCTTGAGCCCAAGATGGAAAGGTGTGTTG 780
QY 1145 CAGTGGATGGGAATGATGTTAGGAGAGCCCGTCGAGGGCCAGTGCACCCAACTACT 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 CAGTGGATGGGAATGATGTTAGGAGAGCCCGTCGAGGGCCAGTGCACCCAACTACT 840
QY 1205 TCTGATGCTCAGCGGTTCCGTCGAGACCCCGCCCAAGCTGGAAAGCCCGCAGATCCGAC 1264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 TCTGATGCTCAGCGGTTCCGTCGAGACCCCGCCCAAGCTGGAAAGCCCGCAGATCCGAC 900
QY 1265 CCGCAGAAAGACCCCTTATGACCCCTGCGCGGCTGTGCTTATGTTACATTGACAGAC 1324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 CCGCAGAAAGACCCCTTATGACCCCTGCGCGGCTGTGCTTATGTTACATTGACAGAC 960
QY 1325 CCTTCTCTCTCCCTAGAGGCTGAGGCTGTCACACGCTCTGTGGGGGTGAGAGAC 1384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 CCTTCTCTCTCCCTAGAGGCTGAGGCTGTCACACGCTCTGTGGGGGTGAGAGAC 1020
QY 1385 T-CCTCTGAGAAAGGTGAGAAAGTGGAGGCTATGCTTTGGGGGACAGGCCAGATGAA 1443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 TCCCTCTGAGAAAGGTGAGAAAGTGGAGGCTATGCTTTGGGGGACAGGCCAGATGAA 1080
QY 1444 GTCTGAGGTCGA-GGAGCCAGCTGGCTGGGGCCAGCTGCTATGTAAGGCTTCTAGTTT 1502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 GTCTGAGGTCGA-GGAGCCAGCTGGCTGGGGCCAGCTGCTATGTAAGGCTTCTAGTTT 1140
QY 1503 GTTCTGAGACCCCGCAGCCGACAGCAACCAATTCCAATTAAGTACATTCCAGCTGAAAA 1562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 GTTCTGAGACCCCGCAGCCGACAGCAACCAATTCCAATTAAGTACATTCCCAAAAAA 1200
```

RESULT 14

AAZ47167 standard; DNA; 1815 BP.

AAZ47167:

28-MAR-2000 (first entry)

Mouse sphingosine kinase 1a DNA.

Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
KMN antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
KM cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis; ds.

Mus musculus.

W09961581-A2.

02-DEC-1999.

25-MAY-1999; 99MO-US11521.

26-MAY-1998; 98US-0086657.

PR 11-AUG-1998; 980S-0096049.

(DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

Spiegel S:

WPT: 2000-072612/06.

P-PSDB: AAY56053.

New sphingosine kinase, used to treat diseases involving abnormal cell proliferation, e.g. cancer -

Example 1; Page 113-114; 116pp: English.
This sequence represents the coding region for the mouse sphingosine kinase 1a. The sequence was isolated by searching a dbEST (expressed

CC sequence tag database) using amino acid sequence results from sequenced
CC tryptic digests of the rat sphingosine kinase. Expression of sphingosine
CC kinase in cells results in formation of sphingosine-1-phosphate, a known
CC second messenger, and confers serum-independent growth. Increases
CC proliferation, and suppresses serum-deprivation or ceramide-induced
CC apoptosis. The sphingosine kinase nucleic acid is used: (a) to increase
CC sphingosine kinase content of cells, specifically for reducing cell death
CC and/or increasing cell proliferation; and (b) to produce transfected
CC cells that are used to screen for agents that inhibit or promote
CC sphingosine kinase activity. Agents that reduce sphingosine kinase
CC activity or expression are used: (i) to reduce cell proliferation,
CC specifically for treating cancer, and (ii) to treat diseases associated
CC with abnormal cell migration or motility, particularly cancer, restenosis
CC or diabetic neuropathy (but also atherosclerosis, stroke and Alzheimer's
CC disease), whereas agents that stimulate sphingosine kinase can be used
CC to treat conditions associated with reduced cell proliferation, e.g.
CC developmental retardation.

SQ Sequence 1815 BP; 382 A; 511 C; 533 G; 389 T; 0 other:

Query Match 54.3%; Score 853.4; DB 21; Length 1815;

Best Local Similarity 73.9%; Pred. No. 3.5e-181;

Matches 1171; Conservative 0; Mismatches 391; Indels 23; Gaps 6;

```
QY 1 GCGCCACAGCCGCGCTCCGACGCGCCGCTGGGACAGCGATAGAGAGTGAAGCAGG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 GCGCCACAGCCGCGCTCCGACGCGCGCCGCTGGGACAGCGATAGAGAGTGAAGCAGG 297
QY 61 AGCGCGCGCGACGCGGAGCGCGCC---ACAGCGCCAGGAGCCGCTGGGACGCGGAGCC 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 AGCGCGCGGTACTCTTAGCAGCGCCGCGGAGCAGCGGTGGCCCTGTACAGCGGAGCC 357
QY 117 GCGGCTCAGAGTTATGATTCACAGCGCGCGCGCGCGCTGCTCCGCGCGCTGCGG 176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 CCGGACCTGGCTATGAAACA---GAATGCCCTCGAGGACTCTCCACAGCGCATGACG 414
QY 177 CGTGTGCTGCTCTGTAACCCGCGCGGAGGCAAGGCAAGCCCTTGACAGCTCTCCGAG 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 415 AGTGTGTGCTCTGTAAACCCCGAGGTGGCAAGGCGCAAGGCTCTGAGCTTCCAGAG 474
QY 237 TCACTGCAAGCCCTTTTGGCTGAGGCTGAAATCTCTTACAGCTGATGCTACTAGCG 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 475 CCGTGTGCAAGCCCTTCTCTGAGGAGGACAGAGATACCTTTAACTGATCTCACGACG 534
QY 297 GCGGAACCAAGCCCGGAGCTGTGCTGTCGAGAGAGCTGGCGCTGGAGCGCTTGGT 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 GAAGAACCAATGCGCAGGAGCTGCTGTGTCGAGAGGAGTTGGTCACTGGAGCCCTGGC 594
QY 357 GGTCAATGCTGAGAGCGGCTGATGACGAGGTGGAACGGGCTGATGAGGCGCTGA 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 595 AGTCATGTCCGGGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 654
QY 417 CTGAGAGACCGCCATCCAGAGCCCTGTGTAGCTCTCCAGAGCGCTGAGCAACCGCT 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 655 CTGAGAGACCGCCATCCAGAGAACCCCTGTGTAGCTCTCCAGAGCGCTGAGCAATGCG 714
QY 477 GCGAGCTTCTTCAACATTATCTGTGATGAGCAGAGTACCACAAATGAGACCTCTGAC 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 715 GCGAGCTTCTGTAACCACTATGCTGGGTACGAGCAGGTACATAATGAGACCTGCTCAT 774
QY 537 CAACGTGACGCTATGCTGTGCGCGCGGCTGTGTACCCATTAACCTGTGTGTGCA 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 CAACGTGACGCTGTGTGTGCGCGCGGCTGTGTACCCATTAACCTGTGTGTGCA 834
QY 597 CAGGCTTGGGGCTGCGCTCTTCTGTGCTCAGCTGAGCTGGGGCTTCAATTGCTGA 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 CACTGTCTTGGGCTGCGGCTCTATCTGTGTGCTCAGCTGTGTGTGCTGTGTGCTGA 894
QY 657 TGTGACCTAGAGAGTGAAGATGCGCGCTGTGGGAGAGATGCGCTTCACTGTGGCAC 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 895 CGTGTGACCTGTGAGATGAGATGACAGCGCGCTTGGGGAGATTCGTTTACAGATGGGAC 954
QY 717 CTTCGTGCGTGTGAGACCGCTGTGGCACTACCGCGCGGCACTGTGCTTACCTGCTTAGG 776
```


•
•
•



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:50:29 ; Search time 14 Seconds
(without alignments)
1137.636 Million cell updates/sec

Title: US-09-937-060A-5
Perfect score: 2016
Sequence: 1 MDPAGPGRGVLPKPCRVLV.....CPEPPSWKPGQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2016	100.0	384	1 SPH1_HUMAN	Q9NYA1 homo sapien
2	892	44.2	617	1 SPH2_MOUSE	Q9J1A7 mus musculu
3	886.5	44.0	654	1 SPH2_HUMAN	Q9N9A0 homo sapien
4	103	5.1	297	1 BMRU_BACSV	P39074 bacillus su
5	100.5	5.0	433	1 Y036_SYNY3	O55452 synechocyst
6	99.5	4.9	1065	1 KDC1_HUMAN	O75912 homo sapien
7	93.5	4.6	791	1 KDC1_MOUSE	Q01383 drosophila
8	93	4.6	712	1 DNLJ_RHOMR	P49421 rhodothermu
9	92.5	4.6	439	1 YSM3_CAEEL	Q10123 caenorhabd
10	92.5	4.6	492	1 GVD2_HALNI	Q9JH22 halobacteri
11	92	4.6	782	1 SMAB_MOUSE	O62179 mus musculu
12	92	4.6	1473	1 NAL1_HUMAN	O9C000 homo sapien
13	89	4.4	414	1 PRS6_CAEEL	P46502 caenorhabd
14	89	4.4	533	1 SERA_RAT	O08651 rattus norv
15	87	4.3	533	1 SERA_HUMAN	O43175 homo sapien
16	87	4.3	572	1 POXB_ECOLI	P07003 escherichia
17	86.5	4.3	2470	1 NMC2_MOUSE	O35516 mus musculu
18	86	4.3	390	1 ARP2_SCHPO	O9UWJ1 schizosacch
19	86	4.3	448	1 5H7_MOUSE	P32304 mus musculu
20	85.5	4.2	662	1 LOX2_PIG	P41659 sus scrofa
21	85	4.2	608	1 HFAFC_CAUCR	Q45978 caulobacter
22	85	4.2	827	1 KDGL_CAEEL	Q03303 caenorhabd
23	84.5	4.2	335	1 G3P2_METAC	P58838 methanosarc
24	84.5	4.2	369	1 XYTM_PSEPU	P21335 pseudomonas
25	84.5	4.2	471	1 HH_DROME	Q02936 drosophila
26	84.5	4.2	492	1 GVD2_HALNI	P33958 halobacteri
27	83.5	4.1	586	1 TUI2_SCHPO	Q9UWJ1 schizosacch
28	83	4.1	448	1 5H7_RAT	P32305 rattus norv
29	83	4.1	520	1 CMCH_NOCILA	O51080 nocardia la
30	83	4.1	1247	1 IRBP_HUMAN	P10745 homo sapien
31	82	4.1	668	1 ACOR_ALCEU	P28614 alceligenes
32	82	4.1	873	1 PODK_CLOSY	P22983 clostridium
33	82	4.1	942	1 KDCT_HUMAN	P52824 homo sapien

34	82	4.1	4568	1 DYHB_CHLRE	Q39565 chlamydomon
35	81.5	4.0	334	1 G3P_PYRAB	Q9Y1P1 pyrococcus
36	81.5	4.0	392	1 RURE_ACICA	P42454 acinetobact
37	81.5	4.0	554	1 CXIA_PABOE	P08305 paracoccus
38	81	4.0	425	1 BIDA_SERMA	P36568 serralia ma
39	81	4.0	521	1 CMCH_STRCL	O85728 streptomyce
40	81	4.0	771	1 CTPG_MYCVU	Q10866 mycobacteri
41	81	4.0	1434	1 V665_HSVII	Q00106 ictalurid h
42	80.5	4.0	311	1 V665_HSVII	P03030 escherichia
43	80.5	4.0	801	1 DHGA_ECOLI	P05465 acinetobact
44	80.5	4.0	978	1 PEX6_RAT	P54777 rattus norv
45	80.5	4.0	1047	1 EF3_SCHPO	O94489 schizosacch

ALIGNMENTS

RESULT 1	ID	SPH1_HUMAN	STANDARD:	PRT:	384 AA.
AC	Q9NYA1	Q9NYA13	Q9HD92	Q9NY70	
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Sphingosine kinase 1 (EC 2.7.1.-) (SK 1) (SPK 1).				
GN	SPHK1 OR SPHK OR SPK				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20323213; PubMed=10863092;				
RA	Melander A.J., Carlos-Dias E., Gosink M., Allen J.M., Takacs L.;				
RT	"Human sphingosine kinase: molecular cloning, functional				
RT	characterization and tissue distribution.";				
RL	Gene 251:19-26(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.				
RX	MEDLINE=20263733; PubMed=10802064;				
RA	Nava V.E., Lacana E., Poulton S., Liu H., Sugitara M., Kono K.;				
RA	Milstien S., Kohana T., Spiegel S.;				
RT	"Functional characterization of human sphingosine kinase-1.";				
RT	FEBS Lett. 473:81-84(2000).				
RN	[3]				
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.				
RX	MEDLINE=20407120; PubMed=10947957;				
RA	Pitson S.M., D'Andrea R.J., Vandeleur B., Moretti P.A.B., Xia P.;				
RA	Gamble J.R., Vadas M.A., Waltenberg B.W.;				
RT	"Human sphingosine kinase: purification, molecular cloning and				
RT	characterization of the native and recombinant enzymes.";				
RL	Biochem. J. 350:429-441(2000).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Van Veldhoven P.P., Gijbers S.;				
RT	Submitted (Aug-1999) to the EMBL/Genbank/DDJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Ovary, and Mammary gland;				
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.;				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.;				
RA	Wagatsuna M., Hosoi T., Kaku Y., Kodaira K., Kondo H., Sugawara M.;				
RA	Takatsuki M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.;				
RA	Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saiko K.;				
RA	Yamamoto K., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.;				
RA	Ninomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RT	Submitted (Aug-2000) to the EMBL/Genbank/DDJ databases.				
CC	- FUNCTION: Catalyzes the phosphorylation of sphingosine to form				
CC	sphingosine 1-phosphate (SPP), a lipid mediator with both intra-				
CC	and extracellular functions. Also acts on D-erythro-sphingosine				
CC	and to a lesser extent sphingosine, but not other lipids, such as				
CC	D,L-threo-dihydrosphingosine, N,N-dimethylsphingosine,				

CC diacylglycerol, ceramide, or phosphatidylinositol.
 CC -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate +
 CC ADP.
 CC -1- SUBUNIT: Binds to calmodulin.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in adult
 CC liver, kidney, heart, and skeletal muscle.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF266756; AAF73470.1; -
 CC EMBL: AF238083; AAF73423.1; -
 CC EMBL: AF200328; AAG01980.1; -
 CC EMBL: AK023393; BAB14558.1; -
 CC EMBL: AK022402; BAB14028.1; -
 CC EMBL: AJ245504; CAB92131.1; -
 CC Genew: HGNC:11240; SPHK1.
 CC MIM: 603730; -
 CC InterPro: IPR001206; DAGKC.
 CC InterPro: IPR003622; DAG_Kin_cat.
 CC Pfam: PF00781; DAGKC; 2.
 CC Prodom: PD005043; DAG_Kin_cat; 1.
 CC SMART: SM00046; DAGKC; 1.
 CC Kinase: Kinase; ATP-binding; Calmodulin-binding.
 CC TRANSFERASE: Kinase; ATP-binding; Calmodulin-binding.
 CC FT CONFILCT 6 15 MISSING (IN REF. 4).
 CC FT CONFILCT 11 15 LPRPC -> AKL (IN REF. 4).
 CC FT CONFILCT 251 251 V -> KP (IN REF. 4).
 CC FT CONFILCT 260 260 V -> I (IN REF. 2).
 CC FT CONFILCT 302 302 L -> F (IN REF. 2).
 CC FT CONFILCT 325 325 V -> G (IN REF. 4).
 CC FT CONFILCT 337 337 V -> M (IN REF. 3).
 CC FT SEQUENCE 384 AA: 42517 MW: EB04A7E2034C2DB0 CRC64;
 CC
 CC Query Match 100.0%; Score 2016; DB 1; Length 384;
 CC Best Local Similarity 100.0%; Pred. No. 3.4e-168;
 CC Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 MDPAGGPRGVLPRLPRCVLYLLNPRGKGLALQLFRSHVOPPLAFAEISFTLMLTERRNA 60
 CC 1 MDPAGGPRGVLPRLPRCVLYLLNPRGKGLALQLFRSHVOPPLAFAEISFTLMLTERRNA 60
 CC 61 RELVRSSEELGRMDALVWMSGCLMHEVYVNGIMERPWEAIOKPLCSIPASGNAALASL 120
 CC 61 RELVRSSEELGRMDALVWMSGCLMHEVYVNGIMERPWEAIOKPLCSIPASGNAALASL 120
 CC 61 RELVRSSEELGRMDALVWMSGCLMHEVYVNGIMERPWEAIOKPLCSIPASGNAALASL 120
 CC 121 NHVAGYEYVNTNEDLLTCTLLCRLLSPNNLLSTFASGLFSEVLSLANGFADVDLE 180
 CC 121 NHVAGYEYVNTNEDLLTCTLLCRLLSPNNLLSTFASGLFSEVLSLANGFADVDLE 180
 CC 121 NHVAGYEYVNTNEDLLTCTLLCRLLSPNNLLSTFASGLFSEVLSLANGFADVDLE 180
 CC 181 SEYRRRLGEMRTTCTFLRLAALRTYGRALAYLPVGRVSKTPASPVVVOGGPDAHLVP 240
 CC 181 SEYRRRLGEMRTTCTFLRLAALRTYGRALAYLPVGRVSKTPASPVVVOGGPDAHLVP 240
 CC 181 SEYRRRLGEMRTTCTFLRLAALRTYGRALAYLPVGRVSKTPASPVVVOGGPDAHLVP 240
 CC 241 LEBPVSHMTVVPDDEFLVYLLALSHLSEMPAFAFGACAGVNHILFYVAVGSRAML 300
 CC 241 LEBPVSHMTVVPDDEFLVYLLALSHLSEMPAFAFGACAGVNHILFYVAVGSRAML 300
 CC 241 LEBPVSHMTVVPDDEFLVYLLALSHLSEMPAFAFGACAGVNHILFYVAVGSRAML 300
 CC 301 RLFLAMEGRHMEYECPLVYVAVFRLPRKDGKGVFVDELWVSAVVOGVHNPYFM 360
 CC 301 RLFLAMEGRHMEYECPLVYVAVFRLPRKDGKGVFVDELWVSAVVOGVHNPYFM 360
 CC 301 RLFLAMEGRHMEYECPLVYVAVFRLPRKDGKGVFVDELWVSAVVOGVHNPYFM 360
 CC 361 MVSQVPEPPPMKPOOMPPEEP 384
 CC 361 MVSQVPEPPPMKPOOMPPEEP 384
 CC 361 MVSQVPEPPPMKPOOMPPEEP 384

RESULT 2
 ID SPH2_MOUSE STANDARD: PRT: 617 AA.
 AC 09JUN97; GDB66; 091YAB;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Sphingosine kinase 2 (EC 2.7.1.-) (SK 2) (SPK 2).
 GN SPHK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN-ICR: TISSUE-Brain;
 RX MEDLINE=20347850; PubMed=10751414;
 RA Liu H., Sugita M., Nava V.E., Edsall L.C., Kono K., Poulton S.,
 RA Miletan S., Kohama T., Spiegel S.;
 RT "Molecular cloning and functional characterization of a novel
 RT mammalian sphingosine kinase type 2 isoform."
 RL J. Biol. Chem. 275:19513-19520(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Thompson D., Pyne S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawachi T., Shinagawa A., Shibata K., Yoshino M., Ichi Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada T.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochi H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Bottelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast tumor;
 RA Strausberg R.;
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
 CC sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
 CC and extracellular functions. Also acts on D-erythro-
 CC dihydrosphingosine, D-erythro-sphingosine and L-threo-
 CC dihydrosphingosine.
 CC -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate +
 CC ADP.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF245448; AAF74125.1; -

```

DR EMBL: AF415214; AAL07500.1; -
DR EMBL: AK004951; BAB23694.1; -
DR EMBL: BC006941; AAH06941.1; -
DR MGI: 1861380; SpHx2.
DR InterPro: IPR001206; SpHx2.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKc; 1.
DR Transferase: Kinase; ATP-binding.
FT CONFLICT 252 252 N -> S (IN REF. 1).
FT CONFLICT 510 510 P -> T (IN REF. 1).
FT CONFLICT 548 548 L -> F (IN REF. 1).
SQ SEQUENCE 617 AA; 65618 MM; 40BE2C2C28BBE26A CRC64;

Query Match 44.2%; Score 892; DB 1; Length 617;
Best Local Similarity 39.9%; Pred. No. 5.7e-70;
Matches 192; Conservative 60; Mismatches 109; Indels 120; Gaps 5;

QY 10 VLPRPCRVLTILNPRGKGAQLQIFRSHVOPDLAEAEISFTLMTERRNHAREVSEEL 69
DB 141 LLPRKPRILLIVNPRGGGLAMQRCMDHVVPMISAGLSFNLTQTERQNHARELVQGLST 200
QY 70 GRMDALVMSGDGLMHEVYNGLMERPDETALQKPLCSLPAGSGNALASLNHYAGYEQV 129
DB 201 SEMEGIVVSGDGLYEVLINGLLDRPDWEDAVRMPIGVLPCCSGNALAGAVNHGCFEQV 260
QY 130 TNEDLLTCTLLICRLLSPHLLSHTASGLRFLFSVLSLNGFIADVDLSEKXPRICE 189
DB 261 VGVDDLKMSLLHCGSGHPDLDSVTLASGRCFSFLSVAMGFLSDVDHSERFALGS 320
QY 190 MRFTLGFELRLAALRTYGRGLAYLPV----- 215
DB 321 ARFTLGAVLGASLHTYGRGLSYLPATEPALPIGHSLPRAKSELYLAPAPAAATHSP 380
QY 216 -----GRVGS----- 220
DB 381 LHRVSVDLPRLPOLAVSPGSPEDLPDLSLNGSGPELTGMDGAGADAPLSPDPLPSSP 440
QY 221 -----KTPAS-----PVYVQDGPVDAHLYPRLEPPVSHHTVVPD 254
DB 441 NALKTAOLSPFAEGPPEPPASSGFLPPHSAPEASTWGVDDLPRPLSPDLQDVNTLEG 500
QY 255 EDFVLIALLSHLSGSEMFAPMGCAAGVNHLYFVAGVSRAMLRLFLMEKGRHMEY 314
DB 501 E-FVYMLGILRSHLCADLMAAPHARFDGCVYHLCVRSIGISRAALLRLILAEHGNHSL 559
QY 315 ECPYLVYVVAFAFLEPPDKGVFAVDGLMSEAVNGCVHPNRYMWYSGCYEPPPSMKP 374
DB 560 GCPHIGVAAARAFLEPLTPKGLLTVDGLVEYGPIDQAVHPGLATLTLTG----PAGOKP 615
QY 375 Q 375
DB 616 Q 616

```

```

RA Milstien S., Kohama T., Spiegel S.;
RT "Molecular cloning and functional characterization of a novel
RT mammalian sphingosine kinase type 2 isoform.";
RL J. Biol. Chem. 275:19513-19520(2000).
(2)
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE-Brain;
RX MEDLINE-21154917; PubMed-11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgorge W., Boecher M., Bloecher H., Bauersachs S., Blum H.,
RA Lauder J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Oltenevelde B., Obermaier B., Tanpe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
(3)
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE-Eye, and Lymph;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
(4)
RN SEQUENCE OF 1-354 FROM N.A. (ISOFORM 3).
RC TISSUE-Carcinoma;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isgai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
CC -I- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
CC sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
CC and extracellular functions. Also acts on D-erythro-
CC dihydrosphingosine, D-erythro-sphingosine and L-threo-
CC dihydrosphingosine.
CC -I- CATALYTIC ACTIVITY: Sphingosine + ATP = sphingosine 1-phosphate +
CC ADP.
CC -I- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be
CC produced by alternative splicing.
CC -I- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF245447; AAF74124.1; -
DR EMBL: AL136701; CAB66636.1; -
DR EMBL: BC006161; AAH06161.1; -
DR EMBL: BC010671; AAH10671.1; -
DR EMBL: AK000599; BAA91280.1; -
DR MIM: 607092; -
DR Genew: HGNC:18859; SPHK2.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKc; 1.
DR Transferase: Kinase; ATP-binding.
FT CONFLICT 1 36 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 1 36 FEPAAGLDLILGSLICRGGGSLFLKININVFICKKKK
FT VARSPLIC 292 390 LSVANGTSDVDIOSERRRALSARFTGLTGLATLTTR
FT GRSLTPATVPEASETP -> PRESDSSSTSSACPLMTTA
FT RSCPRAAASMPGSCPLAQOLALGFRSRIQDVNGCGGRIG
FT SLTCGATORTLPAVAREGGGSLFLKININVFICKKKK
FT (IN ISOFORM 3).
FT P -> S (IN REF. 2).
SQ SEQUENCE 654 AA; 69217 MM; F73FCEGC30DA50F CRC64;

Query Match 44.0%; Score 886.5; DB 1; Length 654;
Best Local Similarity 39.8%; Pred. No. 1.9e-69;

```


Db 160 ISYF-----TSALRTV-----SSAASFPMTLK---IDGFEI-----187
QY 250 TVPDEDFVLVALLHSHLGSEMFAPMGRCAGVHLEFYRAGVSRMLRLFLAMEKG 309
Db 188 ----KEEAVMLVNGVOTIGNRIPLPDASTEDGLDLCRN--TNTJALRELMSEOG 241
QY 310 RHMEYECPLYVYVAVFRLEPKDKGFVAVDGLMSEAVQGVHPYFMV 362
Db 242 -SIDRFGELSYVQASRIEITDPAKKA-DMDGEVYTHTPAVIQVLPHIML 292
RESULT 5
Y036.SYNY3
ID Y036.SYNY3 STANDARD: PRT: 433 AA.
AC 055452:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein sl10036.
GN sl10036.
OS Synechocystis sp. (strain PCC 6803)
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.,
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC METHYLGLYOXAL SYNTHASE FAMILY.
CC -1- SIMILARITY: TO B.SOBITILIS BMRU AND TO E.COLI YEGS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64006; BAA10795.1; -
DR HSP: P37066; B93
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR InterPro: IPR004362; DAG_Like.
DR InterPro: IPR004363; methylgl_synth.
DR Pfam: PF00781; DAGC; 1.
DR Pfam: PF02142; MGS; 1.
DR ProDom: PD005043; Dag_Kin_cat; 1.
DR ProDom: PD010761; Methylgl_synth; 1.
DR SMART: SM00046; DAGC; 1.
DR TIGRFAMs: TIGR00160; MGS; 1.
DR PROSITE: PS01335; METHYLGLYOXAL SYNTH; 1.
DR Hypothetical protein: Complete proteome.
KW DOMAIN 1 126 METHYLGLYOXAL SYNTHASE.
FT ACT_SITE 62 62
SQ SEQUENCE 433 AA; 45909 MW; BDB3D47B2990F6C5 CRC64;
Query Match 5.0%; Score 100.5; DB 1; Length 433;
Best Local Similarity 26.8%; Pred. No. 0.34; Indels 23; Gaps 8;
Matches 40; Conservative 26; Mismatches 60;
QY 19 VLNRGCKG---KAIQVFRSHVQPLAEAFSTLMTERRHARELVR-----SEEL 69
Db 129 LIPNVAGQGVNERLDIKELHGEI-NKLTFFSAENVYDQAKELVKKIKQANEQSD 187
QY 70 GRMDALVMS-GDGLMHEVNGLMERPMEVTAIQKPLCSLPAGSGNALASL---NHAG 125
Db 188 GEGDSFIASGDGTGVSVAALV-----NIGI--PLGIIPRGTAANAFSVALGIPTQIPG 240

QY 126 YEOVTNEDLTNCTLLICRRLSPMNILS 154
Db 241 ACOTINRGITRVDTALCNDI--PALLIA 267
RESULT 6
KDGL_HUMAN
ID KDGL_HUMAN STANDARD: PRT: 1065 AA.
AC 075912; Q9NZ49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diacylglycerol kinase, Iota (EC 2.7.1.107) (Diacylceride kinase) (DGK-
DE Iota) (DAG kinase Iota).
GN DGKI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=99047655; PubMed=9830018;
RA Ding L., Traer E., McIntyre T.M., Zimmerman G.A., Prescott S.M.;
RT "The cloning and characterization of a novel human diacylglycerol
RT kinase, DGK-Iota.";
RL J. Biol. Chem. 273:32746-32752(1998).
RN [2]
RP SEQUENCE OF 135-1065 FROM N.A., AND VARIANT PHE-153.
RX MEDLINE=20173854; PubMed=10706894;
RA Bowne S.J., Sullivan L.S., Ding L., Traer E., Prescott S.M.,
RA Birch D.G., Kennan A., Humphries P., Daiger S.P.;
RT "Evaluation of human diacylglycerol kinase Iota, DGKI, a homolog of
RT Drosophila rfga, in inherited retinopathy mapping to 7q.";
RL Mol. Vision 6:6-9(2000).
CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol -> ADP + 1,2-
CC diacylglycerol 3-phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MARCKS HOMOLOG REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF061936; AAC62010.1; -
DR EMBL: AF219939; AAF43006.1; -
DR EMBL: AF219907; AAF43006.1; JOINED.
DR EMBL: AF219908; AAF43006.1; JOINED.
DR EMBL: AF219910; AAF43006.1; JOINED.
DR EMBL: AF219911; AAF43006.1; JOINED.
DR EMBL: AF219912; AAF43006.1; JOINED.
DR EMBL: AF219913; AAF43006.1; JOINED.
DR EMBL: AF219914; AAF43006.1; JOINED.
DR EMBL: AF219915; AAF43006.1; JOINED.
DR EMBL: AF219916; AAF43006.1; JOINED.
DR EMBL: AF219917; AAF43006.1; JOINED.
DR EMBL: AF219918; AAF43006.1; JOINED.
DR EMBL: AF219919; AAF43006.1; JOINED.
DR EMBL: AF219920; AAF43006.1; JOINED.
DR EMBL: AF219921; AAF43006.1; JOINED.
DR EMBL: AF219922; AAF43006.1; JOINED.
DR EMBL: AF219923; AAF43006.1; JOINED.


```

CC CC -1- FUNCTION: UPON CELL STIMULATION CONVERTS THE SECOND MESSENGER
CC DIACYLGLYCEROL INTO PHOSPHATIDATE, INITIATING THE RESYNTHESIS
CC OF PHOSPHATIDYLINOSITOLS AND ATTENUATING PROTEIN KINASE C
CC ACTIVITY (BY SIMILARITY)
CC CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
CC diacylglycerol 3-phosphate.
CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NERVOUS SYSTEM AND MUSCLE.
CC -1- DEVELOPMENTAL STAGE: DGK IS TRANSCRIBED IN THE EMBRYONIC, PUPAL,
CC AND ADULT STAGES, WITH LITTLE EXPRESSION DURING THE LARVAL
CC STAGES. EXPRESSION IN LATE EMBRYOS IS SPECIFIC TO THE CENTRAL
CC NERVOUS SYSTEM AND HEAD.
CC CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11120; BAA01894.1; -
DR EMBL: AE003839; AAF59180.1; -
DR EMBL: X67335; CAA47750.1; -
DR PIR: S25099; S25099.
DR PIR: S28229; S28229.
DR PIR: A46140; A46140.
DR FlyBase: FBgn0004568; Dgk.
DR InterPro: IPR000756; DAGK.
DR InterPro: IPR001206; DAGK.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00609; DAGK; 1.
DR Pfam: PF00781; DAGK; 1.
DR ProDom: PD002939; DAGK; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00045; DAGK; 1.
DR SMART: SM00046; DAGK; 1.
DR KEGG: Transferase; Kinase; 1.
DR FT: DOMAIN 311 372 THR-RICH.
DR FT: DOMAIN 401 405 POLY-ALA.
DR FT: DOMAIN 430 453 GLN-RICH.
DR FT: DOMAIN 534 539 POLY-GLN.
DR FT: DOMAIN 130 256 CATALYTIC-A (POTENTIAL).
DR FT: DOMAIN 564 755 CATALYTIC-B (POTENTIAL).
DR FT: CONFLICT 374 374 M -> T (IN REF. 1).
DR FT: CONFLICT 444 444 Q -> QQQQQ (IN REF. 3).
DR FT: SEQUENCE 791 AA; 87297 MW; 9DFD00E280FD64A6 CRC64;
SQ
Query Match 4.6%; Score 93.5; DB 1; Length 791;
Best Local Similarity 23.1%; Pred. No. 3.1;
Matches 49; Conservative 34; Mismatches 86; Indels 43; Gaps 9;
OY 15 CRVLVLLNPRGKGAOLQFRSHVQPLLAELAEISFTMLTERRNH-----ARELVASE 67
DB 131 CPLLIVFVNPKSGRGQGRDILR-----KFOYMLNPROVVDLSKGGPKCLTLFJK 178
OY 68 ELGRWDALVNVSGGLMHEVYVNGILMERPD-WETALQKQLCSIPAGSGNALASLHVYAGY 126
DB 179 DLPFR-RVICGGDG---TWGWLLEAMDSIELASQPAIGVPLCTGNDLARCPLWGGGY 233
OY 127 EQVNVNEDLLTN-----CTLLRCRLISPMNL-----LSLHTASGLRLSVLSLAWGF 173
DB 234 EGEVNPRLKMFRASTYMLDRWSIEVNTNTPHSDMRKRVILHS-----NMOKVIELSOSV 289
OY 174 IADVLESEKVRRLGEMFTLTGFLRLAALNT 205
DB 290 VVDKSL-WERFEIQRKSVATSWGTAAST 320
RESULT 8
DNL1_RHOMR STANDARD; PRT; 712 AA.
ID DNL1_RHOMR

```

```

AC P49421;
DR 01-FEB-1996 (Rel. 33; Created)
DR 01-FEB-1996 (Rel. 33; Last sequence update)
DR 16-OCT-2001 (Rel. 40; Last annotation update)
DE DNA ligase (EC 6.5.1.2) (polydeoxyribonucleotide synthase [NAD+]).
GN Liga or Lig.
OS Rhotodermus marinus (Rhotodermus obamensis).
OC Bacteria; Bacteroidetes; Sphingobacteriales;
OC Chromatiales; Rhotodermus.
OX NCBI_Taxid=29549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R-21;
RX MEDLINE=95369716; Pubmed=7642120;
RA Thorbjarnardottir S.H., Jonsson Z.O., Andreasson O.S.,
RA Kristianson J.K., Eggertsson G., Palsdottir A.;
RT "Cloning and sequence analysis of the DNA ligase-encoding gene of
RT Rhotodermus marinus, and overproduction, purification and
RT characterization of two thermophilic DNA ligases.";
RL Gene 161:1-6(1995).
CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER
CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
CC DAMAGED DNA. THIS ENZYME IS THERMOSTABLE BEING ACTIVE AT 5-75
CC DEGREES CELSIUS WITH APPARENT OPTIMAL ACTIVITY ABOVE 55 DEGREES
CC CELSIUS.
CC -1- CATALYTIC ACTIVITY: NAD(+) + [deoxyribonucleotide](N) +
CC [deoxyribonucleotide](M) -> AMP + nicotinamide nucleotide +
CC [deoxyribonucleotide](N+M).
CC -1- MISCELLANEOUS: THIS ENZYME IS THERMOSTABLE.
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10483; AAA93198.1; -
DR HSSP: 087703; 1B04.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR004150; DNA_ligase_OB.
DR InterPro: IPR001679; DNALigase.
DR InterPro: IPR003265; Endo_3C.
DR InterPro: IPR003583; HNH_1.
DR InterPro: IPR000445; HNH.
DR InterPro: IPR004149; Znf_DNALLigase_C4.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00633; HNH; 1.
DR Pfam: PF01653; DNA_ligase_N; 1.
DR Pfam: PF03119; DNA_ligase_ZBD; 1.
DR Pfam: PF03120; DNA_ligase_OB; 1.
DR ProDom: PD003944; DNALLigase; 1.
DR SMART: SM00292; BRCT; 1.
DR SMART: SM00478; Endo3C; 1.
DR SMART: SM00278; Hnh1; 3.
DR SMART: SM00532; LIGANC; 1.
DR TIGRFAMS: TIGR00575; dnlj; 1.
DR PROSITE: PS01072; BRCT; 1.
DR PROSITE: PS01052; DNA_LIGASE_N1; 1.
DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
DR Ligase; DNA repair; DNA replication; NAD.
DR FT: DOMAIN 633 712 BRCT.
DR FT: BINDING 153 153 AMP (BY SIMILARITY).
DR FT: SEQUENCE 712 AA; 79487 MW; 893AD3A78E77FEFC1 CRC64;
SQ
Query Match 4.6%; Score 93; DB 1; Length 712;
Best Local Similarity 21.8%; Pred. No. 2.9;
Matches 63; Conservative 37; Mismatches 109; Indels 80; Gaps 11;

```

```
OY 59 HARELYRSEBELGMDALVYMSGDLHMEVYNGLM-----ERPDMETAIOKPLCS----- 107
DB 401 HNEDEVRSRDRIRIGDLVYIRAGDVYIPQVYRPVYERGRNGNEP--WRNPERPCSGSOLVR 459
OY 108 LKAGSGNLAAS-----LHNYAGYEQVYNEEDLLNCLLCRR--LISPA----- 150
DB 460 LRGADYLVYASDCRQAPVRLLEHFRAGRDANDIEGSGVAVAROLABSGLYRPSLDYLRK 519
OY 151 --NLSLHTASGLRFLSVSLANGFTADVDESERYRLGEMKREFTLGTFLRLAALRTYNG 208
DB 520 LBDLKLBEFAFTRARNLRA-----IEASKORPLSLRFLGLG-----IKHWGK 563
OY 209 RLAYLPVCGVSKTPASPVVYQ-----GPVDAHLVP----- 240
DB 564 TTAELLYORFASIDELAAVITDEIALLEGVGPITAESIANMFRVEDNRRLIEELKELGYN 623
OY 241 ---LEBPVPSHMTVPDEPVLVALLH---SHLGSEMFAPMCRCAAGV 284
DB 624 TORLPEAPAAESPVRGKTFTVLGALPHLTRKEAEELIKRAGGVASSV 672

RESULT 9
ID YSM3_CAEEL STANDARD: PRT: 439 AA.
AC 010123:
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein f52c9.3 in chromosome III.
GN f52c9.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Pelioderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RC SEQUENCE FROM N.A.
RA STRAIN-Bristol N2;
RA Favalio T.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN 12
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U39850; AA81050.2;
DB WormPep: f52c9.3; CE29806;
DR InterPro: IPR001206; DACC;
DR InterPro: IPR003622; DAG_Kin_cat;
DR ProDom: PD005043; DAG_Kin_cat; I;
DR SMART: SM00046; DAGKc; I;
KW Hypothetical protein.
SQ
SEQUENCE 439 AA: 49974 MW: 779850DCG9662DBB CRC64:

Query Match 4.6%; Score 92.5; DB 1; Length 439;
Best Local Similarity 26.5%; Pred. No. 1.7;
Matches 27; Conservative 14; Mismatches 58; Indels 3; Gaps 1;
```

```
RESULT 10
ID GV22_HALNI STANDARD: PRT: 492 AA.
AC 09HHT2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GVPD Protein 2.
GN GVPD2 OR VNG6240G.
OS Halobacterium sp. (strain NRC-1).
OC Plasmid pNRC200.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=40504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroana J.,
RA Swartzell S., Welt D., Hall J., Dahl T.A., Welt R., Goo Y.A.,
RA Iaitenauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Datsarova S.;
RA "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000)
CC
CC -i- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC ASSEMBLY OF GAS VESTICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC REGULATORY FUNCTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF005156; AAG20894.1;
DB Gas vesicle; ATP-binding; Plasmid; Complete proteome.
KW NP_BIND; 39
FT
SQ
SEQUENCE 492 AA: 53928 MW: 655965FDIC36331B CRC64:

Query Match 4.6%; Score 92.5; DB 1; Length 492;
Best Local Similarity 24.3%; Pred. No. 2;
Matches 89; Conservative 36; Mismatches 146; Indels 95; Gaps 16;
```


